

日 本 国 特 許 庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office

出 願 年 月 日

Date of Application:

2003年 2月27日

出 願 番 号

Application Number:

特願2003-050244

[ST.10/C]:

[JP2003-050244]

出 願 人

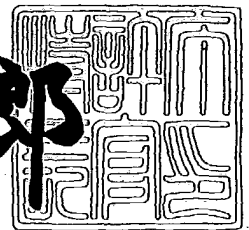
Applicant(s):

株式会社日立製作所

2003年 4月25日

特 許 庁 長 官
Commissioner,
Japan Patent Office

太田 信一郎



出証番号 出証特2003-3030737

【書類名】 特許願

【整理番号】 K02017501A

【あて先】 特許庁長官殿

【国際特許分類】 G06F 3/06

【発明者】

【住所又は居所】 神奈川県小田原市中里 3 2 2 番地 2 号 株式会社日立製作所 RAIDシステム事業部内

【氏名】 武田 貴彦

【発明者】

【住所又は居所】 神奈川県小田原市中里 3 2 2 番地 2 号 株式会社日立製作所 RAIDシステム事業部内

【氏名】 安積 義弘

【発明者】

【住所又は居所】 神奈川県小田原市中里 3 2 2 番地 2 号 株式会社日立製作所 RAIDシステム事業部内

【氏名】 山神 憲司

【発明者】

【住所又は居所】 神奈川県小田原市中里 3 2 2 番地 2 号 株式会社日立製作所 RAIDシステム事業部内

【氏名】 鈴木 勝喜

【発明者】

【住所又は居所】 神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所 システム開発研究所内

【氏名】 白銀 哲也

【特許出願人】

【識別番号】 000005108

【氏名又は名称】 株式会社日立製作所

【代理人】

【識別番号】 100075096

【弁理士】

【氏名又は名称】 作田 康夫

【手数料の表示】

【予納台帳番号】 013088

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 データ処理システム

【特許請求の範囲】

【請求項 1】

第一の計算機及び前記第一の計算機と接続される第一のストレージサブシステムを有する第一のストレージシステムと、

第二の計算機及び前記第二の計算機と接続される第二のストレージサブシステムを有する第二のストレージシステムとを有し、

前記第一及び第二の計算機は第一の通信線で接続され、

前記第一のストレージサブシステムと前記第二のストレージサブシステムは第二の通信線で接続されており、

前記第一のストレージサブシステムは、前記第一の計算機の指示に従って、前記第一のストレージサブシステムが有する第一の記憶領域に対するデータの更新をジャーナルとして第二の記憶領域に記録し、

前記第一の計算機は、前記第一の通信線を介して前記第二の計算機に前記第一のストレージサブシステムにおけるジャーナルの蓄積に関する情報を送信し、

前記第二の計算機は、受信した前記ジャーナルの蓄積に関する情報に基づいて、前記第二のストレージサブシステムに対するコマンドを発行し、

前記第二のストレージサブシステムは、前記コマンドに応答して、前記第一のストレージサブシステムに対して、前記ジャーナルの送付を要求するコマンドを送信し、

前記第一のストレージサブシステムは、前記第二のストレージサブシステムからのコマンドに応答して、要求されたジャーナルを前記第二の通信線を介して転送し、

前記第二のストレージサブシステムは、前記転送されたジャーナルを該第二のストレージサブシステムが有する第三の記憶領域に格納することを特徴とするデータ処理システム。

【請求項 2】

請求項 1 記載のデータ処理システムであって、

前記第二のストレージサブシステムは、前記第一のストレージサブシステムの
前記第一の記憶領域に対応する第四の記憶領域を有し、

前記第二のストレージサブシステムは、前記第二の計算機の指示に従って、前
記第一の記憶領域に格納されているデータの転送を要求する第二のコマンドを前
記第一のストレージサブシステムに送信し、

前記第二のコマンドに従って送信されたデータを前記第四の記憶領域に格納す
ることを特徴とするデータ処理システム。

【請求項3】

請求項2記載のデータ処理システムであって、

前記第二のストレージサブシステムは、前記第二の計算機の指示に従って、前
記第三の記憶領域に格納されたジャーナルを用いて前記第四の記憶領域に格納さ
れたデータを更新することを特徴とするデータ処理システム。

【請求項4】

請求項1記載のデータ処理システムにおいて、

前記第二のストレージシステムが発行する前記第二のストレージサブシステム
に対するコマンドには、

前記第一のストレージサブシステムを指定する情報、前記第二のストレージサ
ブシステムを指定する情報、前記第二及び第三の記憶領域を指定する情報、前記
第二及び第三の記憶領域の部分を指定する情報が含まれていることを特徴とする
データ処理システム。

【請求項5】

請求項1記載のデータ処理システムにおいて、前記第二の計算機が前記第二の
ストレージサブシステムへ発行する前記コマンドの発行のタイミングは、ユーザ
によって指定される任意のタイミングであることを特徴とするデータ処理システ
ム。

【請求項6】

請求項3記載のデータ処理システムにおいて、

前記第二の計算機の指示は、所定のタイミングで行われることを特徴とするデ
ータ処理システム。

【請求項 7】

請求項 3 記載のデータ処理システムにおいて、

前記ジャーナルには、前記第一の記憶領域を更新した時間に関する情報が含まれ、

前記第二の計算機の指示には、所定の時間に関する情報が含まれ、

前記第二のストレージサブシステムは、前記所定の時間以前に更新された前記第一の記憶領域に関するジャーナルを用いて、前記第四の記憶領域に格納されたデータを更新することを特徴とするデータ処理システム。

【請求項 8】

請求項 2 記載のデータ処理システムにおいて、

前記第一、第二、第三又は第四の記憶領域が、複数の論理ボリュームから構成されていることを特徴とするデータ処理システム。

【請求項 9】

請求項 2 記載のデータ処理システムにおいて、

前記第二の記憶領域は、前記第一の計算機で実行されるユーザアプリケーションプログラムには使用されない記憶領域であることを特徴とするデータ処理システム。

【請求項 10】

請求項 1 記載のデータ処理システムにおいて、前記第二のストレージサブシステムは、前記第一のストレージサブシステムの前記第一の記憶領域に対応する第四の記憶領域を有し、

前記第二のストレージシステムは、前記第三の記憶領域に格納されたジャーナルを読み出し、読み出したジャーナルを用いて、前記第四の記憶領域に格納されたデータを更新することを特徴とするデータ処理システム。

【請求項 11】

第一の計算機及び前記第一の計算機と接続される第一のストレージサブシステムを有する第一のストレージシステムと、

第二の計算機及び前記第二の計算機と接続される第二のストレージサブシステムを有する第二のストレージシステムとを含むデータ処理システムにおいて、

前記第一及び第二の計算機は第一の通信線で接続され、

前記第一のストレージサブシステムと前記第二のストレージサブシステムは第二の通信線で接続されており、

前記第一のストレージサブシステムは、前記第一の計算機の指示に従って、前記第一のストレージサブシステムが有する第一の記憶領域に対するデータの更新をジャーナルとして第二の記憶領域に記録し、

前記第一の計算機は、前記ジャーナルの蓄積に関する情報を前記第一のストレージサブシステムから取得し、

前記第一の計算機は、前記取得された前記ジャーナルの蓄積に関する情報に基づいて、前記第一のストレージサブシステムに対するコマンドを発行し、

前記第一のストレージサブシステムは、前記コマンドに応答して、前記第二のストレージサブシステムに対して、前記ジャーナルを送信し、

前記第二のストレージサブシステムは、前記第二の通信線を介して転送されたジャーナルを該第二のストレージサブシステムが有する第三の記憶領域に格納することを特徴とするデータ処理システム。

【請求項 12】

請求項 11 記載のデータ処理システムにおいて、前記第二のストレージサブシステムは、前記第一のストレージサブシステムの前記第一の記憶領域に対応する第四の記憶領域を有し、

前記第二のストレージシステムは、前記第三の記憶領域に格納されたジャーナルを読み出し、読み出したジャーナルを用いて、前記第四の記憶領域に格納されたデータを更新することを特徴とするデータ処理システム。

【請求項 13】

請求項 1 記載のデータ処理システムにおいて、

さらに第三の計算機及び前記第三の計算機と接続される第三のストレージサブシステムを有する第三のストレージシステムを有し、

前記第一の計算機と前記第三の計算機は前記第一の通信線を介して接続されており、

前記第一のストレージサブシステムと前記第三のストレージサブシステムは第

三の通信線で接続されており、

前記第一の計算機は、前記ジャーナルの蓄積に関する情報を前記第一の通信線を介して前記第三の計算機に送信し、

前記第三の計算機は、受信した前記ジャーナルの蓄積に関する情報に基づいて、前記第三のストレージサブシステムに対するコマンドを発行し、

前記第三のストレージサブシステムは、前記コマンドに応答して、前記第一の記憶装置システムに対して、前記ジャーナルの送付を要求するコマンドを送信し、

前記第一のストレージサブシステムは、前記第三のストレージサブシステムからのコマンドに応答して、要求されたジャーナルを前記第三の通信線を介して転送し、

前記第三のストレージサブシステムは、前記転送されたジャーナルを該第三のストレージサブシステムが有する第五の記憶領域に格納することを特徴とするデータ処理システム。

【請求項 1 4】

第一の計算機及び前記第一の計算機と接続される第一のストレージサブシステムを有する第一のストレージシステムと、

第二の計算機及び前記第二の計算機と接続される第二のストレージサブシステムを有する第二のストレージシステムと、

前記第一の計算機システム及び前記第二の計算機システムに接続される記憶装置とを有し、

前記第一及び第二の計算機は通信線で接続されており、

前記第一のストレージサブシステムは、前記第一の計算機の指示に従って、前記第一のストレージサブシステムが有する第一の記憶領域に対するデータの更新をジャーナルとして第二の記憶領域及び前記記憶装置に記録し、

前記第一の計算機は、前記通信線を介して前記第二の計算機に前記第一のストレージサブシステムにおけるジャーナルの蓄積に関する情報を送信し、

前記第二の計算機は、受信した前記ジャーナルの蓄積に関する情報に基づいて、前記第二のストレージサブシステムに対するコマンドを発行し、

前記第二のストレージサブシステムは、前記コマンドに応答して、前記記憶装置から前記ジャーナルを読み出し、

前記第二のストレージサブシステムは、前記転送されたジャーナルを該第二のストレージサブシステムが有する第三の記憶領域に格納することを特徴とするデータ処理システム。

【請求項15】

請求項14記載のデータ処理システムであって、

前記第二のストレージサブシステムは、前記第一のストレージサブシステムの前記第一の記憶領域に対応する第四の記憶領域を有し、

前記第一のストレージサブシステムは、前記第一の計算機の指示に従って、前記第一の記憶領域に格納されているデータを前記記憶装置へ転送し、

前記第二のストレージサブシステムは、前記第二の計算機の指示に従って、前記第一の記憶領域に格納されているデータの転送を要求する第二のコマンドを前記記憶装置に送信し、

前記第二のコマンドに従って送信されたデータを前記第四の記憶領域に格納することを特徴とするデータ処理システム。

【請求項16】

請求項15記載のデータ処理システムであって、

前記第二のストレージサブシステムは、前記第二の計算機の指示に従って、前記第三の記憶領域に格納されたジャーナルを用いて前記第四の記憶領域に格納されたデータを更新することを特徴とするデータ処理システム。

【請求項17】

第一の計算機及び前記第一の計算機と接続される第一のストレージサブシステムを有する第一のストレージシステムと、

第二の計算機及び前記第二の計算機と接続される第二のストレージサブシステムを有する第二のストレージシステムとを有し、

前記第一及び第二の計算機は第一の通信線で接続され、

前記第一のストレージサブシステムは、さらに、第三の計算機及び第一の複数のストレージサブシステムを有し、

前記第三の計算機は、前記第一の複数のストレージサブシステムが有する記憶領域から第一及び第二の仮想的な記憶領域を作成して前記第一の計算機に提供し

前記第二のストレージサブシステムは、さらに、第四の計算機及び第二の複数のストレージサブシステムを有し、

前記第四の計算機は、前記第二の複数のストレージサブシステムが有する記憶領域から第三及び第四の仮想的な記憶領域を作成して前記第二の計算機に提供し

前記第三の計算機と前記第四の計算機は第二の通信線で接続されており、

前記第一のストレージサブシステムは、前記第一の計算機の指示に従って、前記第一の仮想的記憶領域に対するデータの更新をジャーナルとして前記第二の仮想的な記憶領域に記録し、

前記第一の計算機は、前記第一の通信線を介して前記第二の計算機に前記第一のストレージサブシステムにおけるジャーナルの蓄積に関する情報を送信し、

前記第二の計算機は、受信した前記ジャーナルの蓄積に関する情報に基づいて、前記第二のストレージサブシステムに対するコマンドを発行し、

前記第二のストレージサブシステムは、前記コマンドに応答して、前記第一のストレージサブシステムに対して、前記ジャーナルの送付を要求するコマンドを送信し、

前記第一のストレージサブシステムは、前記第二のストレージサブシステムからのコマンドに応答して、要求されたジャーナルを前記第二の通信線を介して転送し、

前記第二のストレージサブシステムは、前記転送されたジャーナルを前記第三の仮想的な記憶領域に格納することを特徴とするデータ処理システム。

【請求項 1 8】

請求項 1 記載のデータ処理システムにおいて、

さらに第三の計算機及び前記第三の計算機と接続される第三のストレージサブシステムを有する第三のストレージシステムを有し、

前記第一の計算機と前記第三の計算機は前記第一の通信線を介して接続され、

前記第一のストレージサブシステムと前記第三のストレージサブシステムは第三の通信線で接続されており、

前記第二の計算機又は前記第二のストレージサブシステムの障害を検出した前記第一の計算機が、前記第三の計算機に前記第一の通信線を介して前記ジャーナルの蓄積に関する情報を送信し、

前記第三の計算機は、受信した前記ジャーナルの蓄積に関する情報に基づいて、前記第三のストレージサブシステムに対するコマンドを発行し、

前記第三のストレージサブシステムは、前記コマンドに応答して、前記第一のストレージサブシステムに対して、前記ジャーナルの送付を要求するコマンドを送信し、

前記第一のストレージサブシステムは、前記第三のストレージサブシステムからのコマンドに応答して、要求されたジャーナルを前記第三の通信線を介して転送し、

前記第三のストレージサブシステムは、前記転送されたジャーナルを該第三のストレージサブシステムが有する第五の記憶領域に格納することを特徴とするデータ処理システム。

【請求項 19】

請求項 1 記載のデータ処理システムにおいて、

さらに第三の計算機及び前記第三の計算機と接続される第三のデータ処理システムを有する第三のストレージシステムを有し、

前記第二の計算機と前記第三の計算機は、前記第一の通信線を介して接続されており、

前記第一のストレージサブシステムと前記第三のストレージサブシステムは第三の通信線で接続されており、

前記第一の計算機又は前記第一のストレージサブシステムの障害を検出した前記第二の計算機が、前記第三の計算機に前記ジャーナルの蓄積に関する情報を前記第一の通信線を介して送信し、

前記第三の計算機は、受信した前記ジャーナルの蓄積に関する情報に基づいて、前記第三のストレージサブシステムに対するコマンドを発行し、

前記第三のストレージサブシステムは、前記コマンドに応答して、前記第二のストレージサブシステムに対して、前記ジャーナルの送付を要求するコマンドを送信し、

前記第二のストレージサブシステムは、前記第三のストレージサブシステムからのコマンドに応答して、要求されたジャーナルを前記第三の通信線を介して転送し、

前記第三のストレージサブシステムは、前記転送されたジャーナルを該第三のストレージサブシステムが有する第五の記憶領域に格納することを特徴とする計算機システム。

【請求項 2 0】

第一の計算機及び第一のストレージサブシステムを有するサービスプロバイダにおけるサービス提供方法であって、

ユーザが、前記第一の計算機を、前記ユーザが有する第二の計算機に登録し、

前記ユーザは、前記第二の計算機に接続される第二のストレージサブシステムに格納されたデータの複製を格納する第三のストレージサブシステムを有し、

前記第二の計算機は、前記第三のストレージサブシステム又は前記第三のストレージサブシステムに接続される第三の計算機の障害を検出した場合に前記第一の計算機を選択して前記第一の計算機及び前記第一のストレージサブシステムの使用を開始し、

前記サービスプロバイダは、前記第一の計算機及び前記第一のストレージサブシステムの使用の開始に伴って前記ユーザに課金することを特徴とする方法。

【発明の詳細な説明】

【0 0 0 1】

【発明の属する技術分野】

本発明は、第一のストレージシステムに格納されたデータを、第二のストレージシステムに複製するための技術に関する。

【0 0 0 2】

【従来技術】

近年、常に顧客に対して継続したサービスを提供するために、第一のストレージ

ジシステムに障害が発生した場合でもデータ処理システムがサービスを提供できるよう、システム間でのデータの複製に関する技術が重要になっている。第一のストレージシステムに格納された情報を第二のストレージシステムに複製する技術としては、特許文献に開示された技術が存在する。特許文献には、第一のストレージシステムに含まれる計算機（以下、「プライマリホスト」という）が、プライマリホストに接続されたディスクアレイ装置（以下、「プライマリディスクアレイ装置」）に格納されたデータを、通信リンク及び第二のストレージシステムに含まれる計算機（以下、「セカンダリホスト」という）を介して、セカンダリホストに接続されたディスクアレイ装置（以下「セカンダリディスクアレイ装置」）に転送する技術（以下、「従来技術」という）が開示されている。

【0003】

一方、コンピュータネットワークの発達により、企業等の有する情報処理システムがより複雑になっている。このことから、ネットワークに接続される機器を一括して管理したいという要求が高まっている。このような要求を満たす技術として、ファイバチャネルあるいはインターネット等のネットワークを介して接続された複数のストレージサブシステムを一括して管理し、仮想的に一つ又は複数のストレージサブシステムとしてユーザに提供するバーチャリゼーションという技術が考案されている。これにより、情報処理システムを使用するユーザは、複数のストレージサブシステムを、あたかも1つのストレージサブシステムであるかのように使用することができる。

【0004】

以下、「ストレージサブシステム」は、単体のハードディスクドライブ、複数のハードディスクドライブの集合体、若しくは専用の制御部で複数のハードディスクドライブを制御するディスクアレイ装置等の記憶装置システムを示す。

以下、「ストレージシステム」及び「サイト」は、1つまたは複数のホストコンピュータと1つまたは複数のストレージサブシステムが接続された情報処理システムを示す。尚、以下、ホストコンピュータは単にホストと略す場合がある。

【0005】

【特許文献1】

米国特許 5170480 号公報

【0006】

【発明が解決しようとする課題】

ここで、上記従来技術を複雑化した情報処理システムに適用することを考える。

【0007】

従来技術においては、プライマリホストとセカンダリホストの間で、双方のディスクアレイ装置に格納されたデータの転送が行われる。つまり、各ホストがデータ転送の経路として使用される。また、ネットワークに接続された各ホストは、ネットワークに接続される複数のディスクアレイ装置に関する情報（ネットワークアドレス等）を保持している。

【0008】

したがって従来技術では、多数存在するディスクアレイ装置のいずれかを各ホストで適当に選択して、選択されたディスクアレイ装置にデータを容易に複製することが出来る。特に従来技術をバーチャリゼーションに適用する場合、バーチャリゼーションを管理している装置（計算機又はスイッチ）と各ホストとを連携させることができる。しかし、各ディスクアレイ装置に格納されたデータがホスト間の通信リンクを介して転送されるので、各ホストのチャネル負荷及びホスト間を接続する回線のトラフィックが増加するという問題点がある。

【0009】

【課題を解決するための手段】

上記課題を解決するために、本発明では、プライマリ及びセカンダリホストが、各ホスト上で動作するソフトウェアに基づいて、各ホストに接続されたストレージサブシステム、例えば、各々プライマリ又はセカンダリディスクアレイ装置の状態を監視する。また、必要に応じて、プライマリ又はセカンダリホストは、ディスクアレイ装置間のデータ転送をプライマリ又はセカンダリディスクアレイ装置に指示する。さらに、各ホストは、各ディスクアレイ装置間のデータ転送を行うための情報をホスト間通信で遣り取りする。一方、各ディスクアレイ装置に格納されたデータは、ディスクアレイ装置間の専用線を介して転送される。尚、

専用線ではなく、テープ装置等の可搬記憶媒体を用いてデータを転送する形態も考えられる。

【 0 0 1 0 】

また、本発明においては、プライマリディスクアレイ装置は、プライマリディスクアレイ装置に格納されたデータの更新の情報をジャーナル（更新履歴）として格納する。ジャーナルは、具体的には、更新に用いられたデータのコピーとメタデータの記録である。さらに、プライマリディスクアレイ装置は、このジャーナルを、各ホストの指示にしたがって、セカンダリディスクアレイ装置に転送する。セカンダリディスクアレイ装置は、セカンダリホストの指示にしたがって、プライマリディスクアレイ装置から受取ったジャーナルを用いてセカンダリディスクアレイ装置に格納されたデータをプライマリディスクアレイ装置で更新されたのと同じ様に更新する。この更新は、プライマリディスクアレイ装置の更新をセカンダリディスクアレイ装置で再現することから、「リストア」と呼ばれる。

【 0 0 1 1 】

尚、ジャーナルの転送は、セカンダリディスクアレイ装置がプライマリディスクアレイ装置にジャーナルコピー命令を発行することで実現する構成としても良い。

【 0 0 1 2 】

また、各ホストに接続されるディスクアレイ装置は、仮想化を制御する装置によって仮想化されたストレージサブシステムでも良い。この場合、データ転送は、仮想化を制御する装置間で、あるいは、仮想化を制御する装置に接続されたストレージサブシステムと、他の同システムで行われる。

【 0 0 1 3 】

【発明の実施の形態】

図 1 は、本発明を適用したデータ処理システムの第一の実施形態のハードウェア構成を示す図である。

【 0 0 1 4 】

本データ処理システムは、プライマリホスト 1 0 0 A 及びプライマリストレージサブシステム（プライマリディスクアレイ装置） 2 0 0 A を有する第一のスト

レージシステム（以下「プライマリストレージシステム」または「プライマリサイト」）10、セカンダリHOST100B及びセカンダリストレージサブシステム（セカンダリディスクアレイ装置）200Bを有する第二のストレージシステム（以下「セカンダリストレージシステム」または「セカンダリサイト」）20、並びにリモートコンソール40を有する。本実施形態では、ストレージサブシステムとしてディスクアレイ装置を例に説明するが、発明は特にディスクアレイ装置に限られない。プライマリサイトとプライマリディスクアレイ装置は各々第一のサイトと第一のディスクアレイ装置と呼ばれることがある。同様に、セカンダリサイトとセカンダリディスクアレイ装置は各々第二のサイトと第二のディスクアレイ装置と呼ばれることがある。

【0015】

各HOST100（プライマリHOST100A及びセカンダリHOST100B）は、CPU110、主記憶装置120、及び入出力処理装置130を有する計算機である。具体的には、ワークステーション、マイクロコンピュータ又はメインフレームコンピュータ等である。

【0016】

各ディスクアレイ装置200（プライマリディスクアレイ装置200A及びセカンダリディスクアレイ装置200B）は、記憶制御装置210、複数のディスク装置220及びSVP（Service Processor）230を有する。記憶制御装置210は、HOSTアダプタ211、キャッシュメモリ212、ディスクアダプタ213、プロセッサ214、及び制御メモリ215を有する。

【0017】

プライマリHOST100Aはプライマリディスクアレイ装置200Aと、セカンダリHOST100Bはセカンダリディスクアレイ装置200Bと、それぞれファイバチャネル66によって接続される。各HOST100のCPU110および主記憶装置120は、入出力処理装置130及びファイバチャネル66を介してディスクアレイ装置200のHOSTアダプタ211に接続される。

【0018】

また、プライマリディスクアレイ装置200Aとセカンダリディスクアレイ装

置200Bは、ファイバチャネル68を介して互いに接続される。プライマリディスクアレイ装置とセカンダリディスクアレイ装置は同じ部屋の中、または同じビルの中にあっても良いが、安全性のために（両装置が同時に同じ障害にあわないよう）距離を置いても良い。プライマリサイト10とセカンダリサイト20との間の距離が長い、すなわち、ファイバチャネル68のデータ転送可能距離を越える場合には、ファイバチャネル68に加え、エクステンダ装置を介してATMなどの広域回線を経由して、各ディスクアレイ装置200間を接続しても良い。

【0019】

リモートコンソール40も、CPUや主記憶装置を有する計算機である。リモートコンソール40、プライマリホスト100A、セカンダリホスト100B、プライマリディスクアレイ装置200A及びセカンダリディスクアレイ装置200Bは、LAN(Local Area Network)やWAN(Wide Area Network)などのIPネットワーク48を介して相互に接続される。すなわち、データ処理システム50はプライマリストレージシステムとセカンダリストレージシステムを接続する少なくとも2つの通信リンク、すなわちディスクアレイデバイス200Aと200Bを接続するファイバチャネル68と、ホスト100Aと100Bを接続するIPネットワーク48を持つ。本実施形態ではこれらの2つの通信リンクは上に述べたように異なる技術としているが、同じ技術（たとえば共にIPネットワーク）としても良い。

【0020】

図10は、図1のデータ処理システムの論理的な構成を示す図である。

【0021】

各ホスト100では、各ディスクアレイ装置200間のデータ転送を制御するプログラムである非同期コピーマネージャ150がCPU110で実行される。非同期コピーマネージャ150は、主記憶装置120に格納される。非同期コピーマネージャ150を実行する各ホスト100は、ホストに直接入力されるまたはあらかじめスケジューリングされたユーザの要求を受けて、各ディスクアレイ装置200が実行するジャーナル処理（ジャーナルの取得処理、ジャーナルの転送処理、及びジャーナルを用いたデータの復元処理）の管理を行なう。詳細は後

述する。

【0022】

また、ジャーナル処理中は各ホスト100の非同期マネージャ150間でIPネットワーク48を用いて随時通信が行われ、ジャーナル処理に必要な管理情報、具体的には、後述するジャーナル作成状況等が交換される。

【0023】

各ディスクアレイ装置200の記憶制御装置210では、コピープログラム2110及びジャーナルバックアップ・リストアッププログラム2120が、プロセッサ214で実行される。これらのプログラムは制御メモリ215に格納される。ジャーナルバックアップ・リストアッププログラム2120は、ジャーナルバックアッププログラムとジャーナルリストアッププログラムから構成される。又、記憶制御装置210は、コピープログラム2110及びジャーナルバックアップ・リストアッププログラム2120の他に、各ホストからの指示に基づいて、ディスク装置220への入出力処理を実行する。

【0024】

ディスク装置220には、1つあるいは複数の論理記憶領域（ボリューム）が作成または関連付けられる。これら複数の論理ボリュームは、ユーザの指定により、データボリューム領域2210とジャーナルボリューム領域2222として使用される。各ホスト100は、非同期コピーマネージャ150を実行することで、各ディスクアレイ装置200が有するジャーナルバックアップ・リストアッププログラム2120とコピープログラム2110の実行を制御する。尚、各ホスト100では、ユーザが使用するアプリケーションプログラム152や、ディスクアレイ装置制御インターフェースとなるプログラム（以下「RAIDマネージャ」）も、そのホスト100が有するCPU110で実行される。又、非同期コピーマネージャ150とRAIDマネージャはプログラム間通信を用いて相互に情報を遣り取りする。

【0025】

リモートコンソール40では、リモートコンソールストレージナビゲータと称するプログラム42が実行される。リモートコンソール40は、プログラム42

を実行することによって、本実施形態のデータ処理システムの各構成要素、具体的には、各ホスト 1 0 0 や各ディスクアレイ装置 2 0 0 の管理を行なう。プログラム 4 2 は、リモートコンソール 4 0 の主記憶装置に格納される。

【 0 0 2 6 】

尚、ここまでに説明した各プログラムは、コンパクトディスクや光磁気ディスクといった可搬記憶媒体を用いて、あるいは、IPネットワーク 4 8 を介して、他の装置から各装置が有する記憶媒体にインストールされる。

【 0 0 2 7 】

図 2 は、第一の実施形態のデータ処理システムの動作の概略を示すフローチャートである。

【 0 0 2 8 】

ステップ 9 1 0 0 でユーザは、ホスト 1 0 0 あるいはリモートコンソール 4 0 が持つ GUI (graphical user interface) を用いてホスト 1 0 0 (プライマリホスト 1 0 0 A でもセカンダリホスト 1 0 0 B でも良い) にペア生成コマンドを入力する。ペア生成コマンドは、ジャーナル取得の対象となる、プライマリディスクアレイ装置 2 0 0 A が有する第一のボリューム (以下「PVOL」) 2 2 1 2 と、PVOL に格納されるデータの複製先 (ペア) となる、セカンダリディスクアレイ装置 2 0 0 B が有する第二のボリューム (以下「SVOL」) 2 2 1 4 とを関連付けるコマンドである。

【 0 0 2 9 】

ペア生成コマンドを受けたホスト 1 0 0 は、プライマリディスクアレイ装置 2 0 0 A において、指定された PVOL 2 2 1 2 に対応するジャーナルを格納するボリューム (以下「ジャーナルボリューム」) 2 2 2 2 A を割り当て、またセカンダリディスクアレイ装置 2 0 0 B において、指定された SVOL 2 2 1 4 に対応するジャーナルを格納するジャーナルボリューム 2 2 2 2 B を割り当てるよう、各ディスクアレイ装置 2 0 0 を制御する (ステップ 9 1 1 0)。ここで、PVOL 2 2 1 2 とそれに割り当てられたジャーナルボリューム 2 2 2 2 A のペアをジャーナルグループと呼び、SVOL 2 2 1 4 とそれに割り当てられたジャーナルボリューム 2 2 2 2 B のペアもジャーナルグループと呼ぶ。ジャーナルグループは「ジャーナル

ペア」と呼ばれることもある。又、ペア生成コマンドはPVOLのジャーナルペアとSVOLのジャーナルペアも関連づける。この関連づけ（すなわちジャーナルペアのペア）を「デバイスグループ」と呼ぶ。

【0030】

尚、PVOL及びSVOLには、記憶容量を拡張するために、単一のボリュームに限らず複数のボリュームの集合（ボリュームグループ）を割り当てることも出来る。この場合、ユーザはペア生成コマンド入力時に、ボリュームグループを指定する。各ディスクアレイ装置200は、指定されたボリュームグループを仮想的な1つのボリュームとして扱い、単一のボリュームに対するのと同様にジャーナル処理を行なう。又、ジャーナルボリュームも同様にボリュームグループとしても良い。

【0031】

尚、本実施形態においては、プライマリホスト100Aがプライマリディスクアレイ装置200Aを主に制御し、セカンダリホスト100Bがセカンダリディスクアレイ装置200Bを主に制御する。したがって、セカンダリホスト100Bにペア生成コマンドが入力された場合、セカンダリホスト100Bは、ペア生成コマンドに含まれる情報（デバイスグループを指定する情報）のうち、プライマリディスクアレイ装置200Aにて必要な情報を、IPネットワーク48を介してプライマリホスト100Aに転送する。同様に、プライマリホスト100Aにペア生成コマンドが入力された場合は、セカンダリディスクアレイ装置200Bにて必要な情報を、セカンダリホスト100Bに転送する。

【0032】

ジャーナルボリュームの割り当て方法としては、少なくとも次の2つが本実施形態で可能である。

(1) ペア生成コマンド入力時に、ユーザ自身がジャーナルボリュームを指定する。

(2) ホスト100が未使用の論理ボリュームを任意に選択して使用する。たとえば、以下の方法が取られる。まず、予め、各ディスクアレイ装置200において、未使用の論理ボリュームが制御メモリ215内にてジャーナルボリュームブ

ールとして管理される。

【0033】

各ディスクアレイ装置200は、ジャーナルボリュームプール内に登録されている論理ボリュームに関する情報、例えばディスクアレイ装置200A内の物理アドレス、物理ボリュームの容量等を各ホスト100に通知する。ペア生成コマンドが入力されたホスト100は、この通知された情報に基づいて、適当なボリューム（単数でも、複数でも構わない。複数の場合、ホスト100はその複数のボリュームを1つの仮想的ボリュームとして扱う）をジャーナルボリュームとして選択する。

【0034】

尚、ユーザは、ホスト100で実行されている通常のアプリケーションが、ジャーナルボリュームに対し入出力要求を発行できるかどうかを指定出来る。これには、（1）通常の入出力処理に使用されるディスク装置220上に割り当てられたボリュームをジャーナルボリュームとして選択する場合と、（2）ホスト100が通常の入出力処理に使用できないボリュームをジャーナルボリュームとして選択する場合がある。

【0035】

前者の場合、ホスト100の通常のアプリケーションからジャーナルの参照が可能である。また、ディスクアレイ装置200にファイバチャネルで接続された別のホスト100のアプリケーションからジャーナルの参照も可能である。従って、通常のアプリケーションで、ジャーナルに関する統計情報の取得や管理を行なうことが可能であるが、誤ってジャーナルを破壊してしまう可能性もある。

【0036】

一方、後者の場合、非同期コピーマネージャ150を実行するホスト100が各ディスクアレイ装置200にジャーナルコピー・リストプログラム又はコピープログラムを実行させる場合にのみ、ホスト100のジャーナルの参照が許されることとなる。したがって、ホスト100からの通常の入出力処理でジャーナルが破壊されることが無い。

【0037】

その後、ジャーナル処理が、たとえばプライマリストレージシステム10で、実行される(ステップ9120)。ジャーナル処理は、取得処理9122、コピー処理9124及びリストア処理9126を含む。プライマリディスクアレイ装置200Aは、プライマリホスト100Aからジャーナルの取得を指示するコマンド(以下「ジャーナル取得開始コマンド」)を受信したことを契機に、ジャーナルの取得処理9122を開始する。ジャーナル取得開始コマンド受領後、プライマリディスクアレイ装置200Aは、PVOL2212へのデータの各書き込み処理(ステップ9200、9210)の後、ジャーナルデータ及びメタデータをジャーナルボリューム2222Aに格納する(ステップ9220)。尚、ジャーナルデータは書き込み処理による更新後データ(つまり、書き込まれたデータ)のコピーであり、メタデータは、更新データがPVOL2212に書込まれた時刻、更新データの格納アドレス、対応するジャーナルデータのジャーナルデータ領域への格納アドレスやデータ長に関する情報である。ジャーナルはジャーナルデータとそれに対応するメタデータから構成される。

【0038】

ジャーナル取得処理9122が開始された後、各ホスト100は、非同期コピーマネージャ150を実行して、定期的にジャーナルコピー処理9124を制御する。ジャーナルコピー処理9124とは、ディスクアレイ装置200間でジャーナルの転送を行う処理である。ジャーナルコピー処理9124は、プライマリホスト100Aが、プライマリディスクアレイ装置200Aから得たジャーナル作成状況に関する情報(詳細は後述)に基づいてジャーナルをコピーする必要があると決定した時に開始される(たとえば、予め決められた量の情報がプライマリディスクアレイ装置200Aのジャーナルボリューム2222Aに格納された時)。次にプライマリホスト100Aは、セカンダリホスト100Bにリンク48を経由して通知する。これにより、セカンダリホスト100Bがセカンダリディスクアレイ装置200Bにデータ読み出しの要求(以下、「ジャーナルコピー要求コマンド」)を発行して、プライマリディスクアレイ装置200Aからのジャーナル転送を開始させる(ステップ9300)。

【0039】

ジャーナルコピー要求コマンドを受取ったセカンダリディスクアレイ装置 200B は、プライマリディスクアレイ装置 200A に対して、データの読み出し要求を発行する（ステップ 9310）。プライマリディスクアレイ装置 200A は、要求されたデータを、コピープログラム 2110 を実行することで、セカンダリディスクアレイ装置 200B に送信する。ジャーナルコピー処理 9124 の詳細は後述する。

【0040】

一方、ジャーナル取得処理 9122 が開始される前に PVOL 2212 に格納されていたデータは、ジャーナルコピー処理 9124 が開始されてもセカンダリディスクアレイ装置 200B には転送されない。したがって、別途 PVOL 2212 から SVOL 2214 へこれらのデータ（以下「イニシャルデータ」）をコピーする必要がある。そのため、本実施形態においては、イニシャルデータを PVOL 2212 から SVOL 2214 に転送する形成コピー処理が実行される（ステップ 9130）。イニシャルデータは、ホスト 100 の指示に基づいて、PVOL 2212 のボリューム先頭領域から末尾まで順に転送される。本処理も、各ディスクアレイ装置 200 のコピープログラム 2110 に基づいて実行される。

【0041】

形成コピーとジャーナルコピー処理 9124 は、非同期かつ並行に行なわれる。つまり、形成コピーは、ペア生成コマンドに基づいて PVOL 2212 及び SVOL 2214 が指定された後であれば、ジャーナル取得処理 9122 及びジャーナルコピー処理 9124 の実行前でも実行中でも行うことが出来る。ただし、形成コピーが完了しない限り、セカンダリディスクアレイ装置 200B においてリストア処理 9126 を行なっても、SVOL 2214 には PVOL 2212 の内容が完全には反映されない。リストア処理 9126 は、コピー処理 9124 に従いプライマリディスクアレイ装置 200A から受取ったジャーナルを用いて、SVOL 2214 へ、PVOL 2212 のデータの更新を反映することを含む。

【0042】

尚、形成コピーを、セカンダリディスクアレイ装置 200B がプライマリディスクアレイ装置 200A にデータを要求する 1 つまたは複数のリードコマンドを

発行することによって実行して、プライマリディスクアレイ装置200Aの負担を軽減できる。

【0043】

すべてのイニシャルデータがセカンダリディスクアレイ装置200BのSVOL2214へコピーされたら、コピープログラム2110Bは、セカンダリホスト100Bへ形成コピー終了を報告する。それ以後セカンダリサイト20でのデータの正確なりカバリが可能になる。一般に形成コピーはジャーナル取得処理開始以後に開始される。

【0044】

ジャーナル取得処理9122は、ホスト100Aが、プライマリディスクアレイ装置200Aへのジャーナル取得の停止を指示するコマンド（以下「ジャーナル取得停止コマンド」）の発行を指示することで終了する。

【0045】

又、ホスト100Bからのコマンド（以下「ジャーナルリストア要求コマンド」）に応じて（ステップ9400）、セカンダリディスクアレイ装置200Bは、ジャーナルボリューム2222Bに格納されたジャーナルを用いてSVOL2214に格納されたデータをリストアする（ステップ9410）。この処理をジャーナルリストア処理9126と言う。ジャーナルリストア処理9126の詳細は後述する。

【0046】

図3は、本発明の計算機システムの第一の実施形態におけるジャーナル取得9122、ジャーナルコピー処理9124、及びジャーナルリストア処理9126の動作を示す図である。ホスト100A及び100Bでそれぞれ非同期コピーマネージャ150を実行して、これらの処理を制御する。プライマリディスクアレイ装置200Aは、ジャーナルバックアップ・リストアプログラム2120のうちジャーナルバックアッププログラム2122を実行する。ジャーナルバックアッププログラム2122を実行することによって、プライマリディスクアレイ装置200Aは、PVOL2212に書き込まれるデータのコピーをジャーナルデータとして、ジャーナルボリューム2222Aへ格納する。又、プライマリディスク

アレイド装置200Aはジャーナルの一部としてメタデータもジャーナルボリューム2222Aへ格納する。上記のステップが、ジャーナル取得処理9122である。

【0047】

一方、セカンダリディスクアレイド装置200Bは、ジャーナルバックアップ・リストアッププログラム2120のうち、ジャーナルリストアッププログラム2124を実行してジャーナルリストアップ処理9126を行なう。ジャーナルリストアッププログラム2124は、ジャーナルボリューム2222Bに格納されたジャーナルをリストアップして、PVOL2212で更新されたデータをデータボリューム2214へ反映する。

【0048】

以下、図3を用いて、ジャーナル取得9122、ジャーナルコピー処理9124及びジャーナルリストアップ処理9126の手順を説明する。

【0049】

プライマリディスクアレイド装置200Aでデータボリューム2210のジャーナル取得処理9122が開始されると、プライマリディスクアレイド装置200Aは、プライマリホスト100AからPVOL2212への書き込み処理5100に応じて、ジャーナルを作成し、作成したジャーナルをジャーナルボリューム2222Aへ格納する。(ステップ5200)。

【0050】

プライマリホスト100Aは、非同期コピーマネージャ150を実行して特定のコマンド(以下「ジャーナル作成状況取得コマンド」)を発行することで、プライマリディスクアレイド装置200Aから、ジャーナル作成状況に関する情報(たとえばジャーナルボリューム内のジャーナルの容量)を取得する。(ステップ5300)。

【0051】

プライマリホスト100Aが取得したジャーナル作成状況に関する情報は、ホスト間で処理を連携するために、IPネットワーク48を経由してセカンダリホスト100Bへ通知される(ステップ5000)。この情報は、具体的には、ジャ

ーナルボリューム2222Aがセカンダリディスクアレイ装置200Bへコピーできるようになった時にホスト100Aからホスト100Bへ通知する際に使用される。

【0052】

セカンダリホスト100Bは、非同期コピーマネージャ150を実行し、GUIを介したユーザからの指示の入力あるいは予め定められていたスケジュール（たとえばプライマリディスクアレイ装置200Aで一定量のジャーナルがジャーナルボリュームに格納された時、または一定期間ごと）に従い、セカンダリディスクアレイ装置200Bに対し、ジャーナルコピー要求コマンドを発行する（ステップ5400）。

【0053】

ジャーナルコピー要求コマンドには、コピーすべきジャーナル（複数でも良い）、そのジャーナルが格納されているジャーナルボリューム、そのジャーナルボリュームを有するディスクアレイ装置200を指定する情報、及びコピーしたジャーナルが格納されるジャーナルボリュームを指定する情報が含まれている。

【0054】

ジャーナルコピー要求コマンドを受信したセカンダリディスクアレイ装置200Bの記憶制御装置210は、コピープログラムを実行することで、リードコマンドをプライマリディスクアレイ装置200Aに対して発行する（ステップ5500）。このリードコマンドを受取ったプライマリディスクアレイ装置200Aは、リードコマンドで指定されたジャーナルをセカンダリディスクアレイ装置200Bに送信する（ステップ5600）。セカンダリディスクアレイ装置200Bに送信されたジャーナルが格納されていた領域はページ（開放）され、新たなジャーナルの格納に利用される。

【0055】

ジャーナルを受信したセカンダリディスクアレイ装置200Bは、ジャーナルコピー要求コマンドで指定されたジャーナルボリューム2222Bに、受信したジャーナルを格納する。その後、セカンダリホスト100Bは、セカンダリディスクアレイ装置200Bに対して、ジャーナルリストア要求コマンドを発行する

(ステップ5700)。

【0056】

ジャーナルリストア要求コマンドを受信したセカンダリディスクアレイ装置200Bは、ジャーナルリストアプログラム2124を実行して、ジャーナルボリューム2222BからSVOL2214ヘデータのリストアを行なう(ステップ5800)。尚、リストアが終わったジャーナルが格納されていた領域はパージ(開放)され、新たなジャーナルの格納に利用される。

【0057】

尚、非同期コピーマネージャ150を実行するホスト100は、ホストフェイルオーバを実行することが出来る。具体的には、プライマリホスト100Aが何らかの理由で利用不能になり、ジャーナルコピー処理の継続が不可能になった場合、セカンダリホスト100Bがプライマリホスト100Aの機能を併せて実行する。

又、プライマリディスクアレイ装置が、ストレージエリアネットワークで複数のプライマリホストに接続されていても良い。こうした環境でも前記ジャーナル取得や他の処理は、当業者によって実行できるよう容易に修正可能である。

【0058】

図4は、本実施形態で用いられるPVOL2212とジャーナルボリューム2222Aの対応を示す図である。以下、ジャーナルボリューム2222Aを正ジャーナルボリュームと呼び、ジャーナルボリューム2222Bを副ジャーナルボリュームと呼ぶ。双方のデータ構造は基本的に同一である。

【0059】

一般に、PVOL、SVOLおよびジャーナルボリュームは各々予め定められた論理ブロック単位で管理される(たとえば512KB)。論理ブロックの各々には、論理ブロックアドレス(以下「LBA」)が付与されている。

【0060】

正ジャーナルボリュームは、メタデータ領域7100及びジャーナルデータ領域7200を有する。ジャーナルデータ領域7200には、先に説明したジャーナルデータ7210、即ち、ライトコマンドによってPVOLに書き込まれたデータ

5110のコピーが格納される。メタデータ領域7100には、先に説明したメタデータ7110、即ち、PVOL2212の更新が行われた時刻、更新データの格納アドレス7112、対応するジャーナルデータ7210のジャーナルデータ領域の格納アドレス7114及び更新データ長が格納される。

【0061】

各アドレスはLBAで、データ長は論理ブロック数で各々表されても良い。また、データが格納されている場所は、データが格納された領域（ジャーナルデータ領域又はメタデータ領域）のベースアドレス（先頭LBA）との差分（オフセット）で表されても良い。本実施形態において、メタデータのデータ長は一定（例えば64バイト）であるが、ジャーナルデータのデータ長は、ライトコマンドで更新されるデータに依存するので一定ではない。

【0062】

ジャーナルグループ定義時に、各ディスクアレイ装置200は、設定されるジャーナルボリューム2222に対して、メタデータ領域7100及びジャーナルデータ領域7200の設定を行う。具体的には、各領域の先頭LBA及びブロック数が指定される。各ホスト100は、非同期コピーマネージャ150を実行して、設定された領域に関する情報（先頭LBA、ブロック数）を要求するコマンド（ジャーナルグループ構成取得コマンド）をディスクアレイ装置200に発行する。これにより、各ホスト100は、各ディスクアレイ装置200が設定したメタデータ領域7100及びジャーナルデータ領域7200に関する情報を取得することができる。

【0063】

図17は、本実施形態で用いられるジャーナルボリューム2222BとSVOL2214の対応を示す図である。副ジャーナルボリュームは、正ジャーナルボリュームと同様にメタデータ領域7100及びジャーナルデータ領域7200を有する。メタデータ領域7100には、正ジャーナルボリュームのメタデータ領域から転送されたメタデータ7110Bが格納される。ジャーナルデータ領域7200には、正ジャーナルボリューム2222Aのジャーナルデータ領域から転送されたジャーナルデータ7210B（メタデータ7110Bに対応する）が格納さ

れる。

【0064】

メタデータはPVOL2112で行われたデータ更新の情報を持ち、そのアドレス情報7114は対応するジャーナルデータ7210（副ボリュームのジャーナルデータ領域へコピーされる）のアドレスを示す。更に、ジャーナルデータ7210を、副ジャーナルボリューム2222Bのジャーナルデータ領域7200からアドレス7112に対応するSVOL2214のアドレスへコピーすることによって、PVOL2212での更新をSVOL2214へ反映できる。

【0065】

図5は、本実施形態の正ジャーナルボリュームと副ジャーナルボリュームのジャーナルデータ領域を示す図である。

【0066】

正ジャーナルボリュームと副ジャーナルボリュームはそれぞれLBAでアドレスづけられていて、正副ジャーナルボリュームの各々のLBAは1対1に対応付けられている。

【0067】

正ジャーナルボリュームが有するジャーナルデータ領域7200は、ジャーナルデータが格納されているジャーナル格納済み領域2232、2233、及び2234と、ジャーナルデータが格納されていないパージ済み領域2231とに区別される。パージ済み領域は、PVOL2212の新たなジャーナルデータの格納に使用される。

【0068】

副ジャーナルボリュームが有するジャーナルデータ領域7200は、既にSVOLへのリストアに使用されたジャーナルデータが格納されている（又はジャーナルデータが格納されていない）リストア済み領域4231、SVOLへのジャーナルリストアの対象として指定されたジャーナルデータが格納されているリストア中領域4232、ジャーナルリストアの対象となっていないジャーナルデータが格納されているリード済み領域4233、及び正ジャーナルボリュームから転送中のジャーナルデータが格納されるリード中領域4234とに区別される。

【0069】

正ジャーナルボリュームのパージ済み領域2231は、副ジャーナルボリュームのリストア中領域4232又はリストア済み領域4231の一部と対応づけられる。

【0070】

正ジャーナルボリュームのジャーナル格納済み領域は、副ジャーナルボリュームのリード済み領域4233、リード中領域4234又はリストア済み領域4231の一部と対応づけられる。ここで、リード済み領域4233に対応するジャーナル格納済み領域2232は、ジャーナルがセカンダリディスクアレイ装置200Bへ送信された後であるので、パージ可能である。又、リード中領域4234に対応するジャーナル格納済み領域2233に格納されたジャーナルデータは、データ転送の対象となっているために、パージすることが出来ない。ジャーナル格納済み領域2232は、対応するジャーナルが転送された後直ちにパージする必要は無い。パージは定期的に行なっても良いし、プライマリホスト100Aからの指示（「ジャーナルパージコマンド」）に従ってパージしても良い。

【0071】

正副ジャーナルボリュームのジャーナルデータ領域7200が有する各領域は、各領域の境界にある論理ブロックのLBAを示すポインタで各ホスト100に識別される。プライマリホスト100Aがプライマリディスクアレイ装置200Aから取得するジャーナル処理状況に関する情報とは、具体的にはこれらポインタの値である。

【0072】

各ホスト100は、接続されたディスクアレイ装置200にジャーナル作成状況取得コマンドを発行することによって、これらポインタの値をディスクアレイ装置200から取得する。その後、ホスト100は、これらポインタの値を用いて、ジャーナルボリュームのどの領域にジャーナルデータが格納されているかを判断する。尚、これらポインタの値は、制御メモリ215に格納されていても良い。

【0073】

以下、各ポインタについて説明する。尚、図5において、LBAは、図の上から下に向けて番号が割り振られている。したがって、図の一番上のLBAの番号が一番小さい。また、ジャーナルボリュームはサイクリックバッファと同じように、繰り返し使用される。つまり、ジャーナルボリュームの末尾の論理ブロックまで使用したら、先頭の論理ブロックが再度使用される。正副ジャーナルボリュームのいずれでも、ジャーナルは作成された順番にジャーナルボリュームに書き込まれる。まず、正ジャーナルボリュームのポインタについて説明する。

【0074】

ジャーナルアウトLBA 2241は、ジャーナル格納済み領域先頭の論理ブロックに対応するLBAを示すポインタである。このポインタで示される論理ブロックに、ページされていないもっとも古いジャーナルデータが格納されている。プライマリHOST 100A又はセカンダリHOST 100Bは、ジャーナルアウトLBA 2241で示されるLBAに対応する論理ブロックを、転送対象となるジャーナルデータの先頭の論理ブロックとして認識する。

【0075】

ジャーナルインLBA 2242は、ジャーナルデータが格納されている末尾の論理ブロックに隣接する空の論理ブロックに対応するLBA、即ち、次にジャーナル取得処理によってジャーナルデータが格納され始める論理ブロックに対応するLBAを示すポインタである。プライマリHOST 100A又はセカンダリHOST 100Bは、ジャーナルインLBA 2242で示されるLBA以上のLBAを有する論理ブロックがジャーナルデータ格納のために使用可能であると認識する。

【0076】

又、プライマリHOST 100A又はセカンダリHOST 100Bは、ジャーナルアウトLBA 2241で示されるLBAから、ジャーナルインLBA 2242で示されるLBAの直前までの領域に、ジャーナルデータが格納されていると認識する。したがって、ジャーナルアウトLBA=ジャーナルインLBAならば、正ジャーナルボリュームのジャーナルデータ領域には、副ジャーナルボリュームに転送すべきジャーナルデータは含まれていないとプライマリHOST 100A又はセカンダリHOST 100Bは判断する。

【0077】

次に、副ジャーナルボリュームのポインタについて説明する。

【0078】

リストア済みLBA 4241は、リストア処理が完了した論理ブロックのうち、最も大きなLBAを有する論理ブロックを示すポインタである。したがって、リストア済みLBAポインタで示されるLBAより小さなLBAを有する論理ブロックは、新たに正ジャーナルボリュームから転送されたジャーナルデータの格納に使用される。即ち、リストア済みLBA 4241以下のLBAを有する論理ブロックでは、ジャーナルデータがパージされている。

【0079】

尚、副ジャーナルボリュームのパージ処理は、リストア処理終了後、記憶制御装置210が自動的にこなされて良い。又、ジャーナルデータのパージは、実際にジャーナルデータを無意味なデータで上書きすること、又は単にポインタを移動してその領域をライト（上書き）可能とすることでも実現できる。PVOLのジャーナル格納済み領域2232のパージと同様に、副ジャーナルボリュームのパージもリストア処理終了後即座に行なう必要はない。

【0080】

リストア予定LBA 4242は、リストア済みLBA 4241より一つ大きいLBAで示される論理ブロックから、リストア予定LBA 4242で示される論理ブロックまでに格納されているジャーナルデータを用いて、SVOLをリストアするジャーナルリストアコマンドがセカンダリホスト100Bから出されていることを示すポインタである。したがって、リストア予定LBA=リストア済みLBAならば、副ジャーナルボリュームには、リストア対象のジャーナルデータが存在しない。

【0081】

リード済みLBA 4243は、プライマリディスクアレイ装置200Aから受取ったジャーナルデータが格納されている論理ブロックのうち、最も大きいLBAを有する論理ブロックを指すポインタである。言い換えると、本ポインタは、セカンダリディスクアレイ装置200Bがプライマリディスクアレイ装置200Aに対して最後に発行したリードコマンドに基づいて転送されたジャーナルデータの

末尾が格納されている論理ブロックを示す。

【0082】

セカンダリHOST100Bは、リード済みLBA4243ポインタによって、本ポインタで示されるLBAに格納されたジャーナルデータに対応する正ジャーナルボリュームのジャーナルデータまでが、副ジャーナルボリュームに格納されたことを確認する。その確認を行ったセカンダリHOST100Bは、プライマリHOST100Aに対して、リード済みLBA4243についての情報を通知する。プライマリHOST100Aは、通知された情報に基づいて、リード済みLBA4243に対応するジャーナルデータが格納された論理ブロックまでジャーナルデータ領域をパージするよう、プライマリディスクアレイ装置200Aに指示できる。尚、ここでのパージも、ジャーナルアウトLBA2241ポインタの移動で実現されても良い。

【0083】

リード予定LBA4244は、セカンダリHOST100Bがセカンダリディスクアレイ装置200Bに発行した最新のジャーナルコピー要求の対象となったジャーナルデータ領域の末尾の論理ブロックのLBAを示すポインタである。したがって、リード予定LBA=リード済みLBAならば、ジャーナルコピーの対象となるジャーナルデータが存在しない。つまり、各ディスクアレイ装置200が実行中のジャーナルコピー処理は無い。

尚、正副ジャーナルボリュームのメタデータ領域同士も図5と同様の対応関係がある。ジャーナルデータ領域と同様に、メタデータ領域用の各ポインタ（ジャーナルアウトLBA、ジャーナルインLBA、リストア済みLBA、リストア予定LBA、リード済みLBA、リード予定LBA。これらはジャーナルデータ領域用のポインタとは別個のポインタである）で各HOST100および記憶制御装置210から制御される。

【0084】

各HOST100は、非同期コピーマネージャ150を双方で実行することで、各ディスクアレイ装置200におけるジャーナル取得状況を、各ポインタの値を取得することで確認する。たとえば、各HOST100は、正副ジャーナルボリュ

ームについて、ジャーナルペア指定時に決定されたジャーナルボリュームの大きさと、ディスクアレイ装置200から取得したポインタの値の差分とから、ジャーナルがジャーナルボリュームにどの程度、例えば何パーセント蓄積されているか算出する。

【0085】

そして、各ホスト100は、算出結果に基づいて、正ジャーナルボリュームに格納されたジャーナルをどこまでパージするか、正ジャーナルボリュームに格納されたジャーナルのうち、何処までをセカンダリディスクアレイ装置200Bへ転送するか、転送されたジャーナルのうち、どこまでをSVOLにリストアするか、等を各ディスクアレイ装置200に対して指示する。

【0086】

例えば、セカンダリホスト100Aは、正ジャーナルボリュームの容量の50パーセント以上にジャーナルが格納されたことを判断した場合に、セカンダリディスクアレイ装置200Bに対してジャーナルコピー要求を発行するとしても良い。

【0087】

ここで、各ホスト100が各ディスクアレイ装置200に対して行う指示とは、具体的には、ジャーナル作成状況取得コマンドと、ジャーナル処理コマンドである。

【0088】

ジャーナル作成状況取得コマンドは、正ジャーナルボリュームにジャーナルがどの程度蓄積されたかの情報を取得するためにプライマリホスト100Aから発行される場合と、副ジャーナルボリュームのリード処理およびリストア処理がどこまで進んでいるかの情報を取得するためにセカンダリホスト100Bから発行される場合がある。

【0089】

ジャーナル処理コマンドは、ジャーナルのパージをプライマリディスクアレイ装置200Aに実行させるためにプライマリホスト100Aから発行される場合と、ジャーナルコピー処理およびジャーナルリストア処理をセカンダリディスク

アレイド装置200Bに実行させるためにセカンダリホスト100Bから発行される場合とがある。従って、ジャーナルコピー要求コマンド及びジャーナルリストア要求コマンドは、ジャーナル処理コマンドの一種である。

【0090】

尚、正ジャーナルボリュームと副ジャーナルボリュームのLBAは1対1対応としたが、正ジャーナルボリュームのLBAから副ジャーナルボリュームのLBAへの適当なアドレス変換手段を用いれば、副ジャーナルボリュームの領域が正ジャーナルボリュームより広く指定されても良い。

【0091】

図6は、ジャーナル取得処理9122、ジャーナルコピー処理9124並びにリストア処理9126の詳細を示すフローチャートである。

【0092】

プライマリホスト100Aは、あらかじめユーザの要求によってスケジューリングされた間隔、又は所定の間隔で、定期的にプライマリディスクアレイド装置200Aのジャーナル格納済み領域に関する情報をジャーナルアウトLBA及びジャーナルインLBAのポインタを用いて取得し（ステップ6100、6200。図3のステップ5300）、取得した情報をセカンダリホスト100Bへ送る（ステップ6110）。

【0093】

セカンダリホストは、通知されたジャーナル格納済み領域を示す情報に基づいて、ジャーナルコピー処理の対象となる正ジャーナルボリュームの論理ブロック領域を決定する。尚、ジャーナルコピーの対象となる正ジャーナルボリュームの論理ブロック領域は、プライマリホスト100Aが予め決定することもできる。

【0094】

その後、セカンダリホスト100Bは、セカンダリディスクアレイド装置200Bに、決定された論理ブロック領域を示す情報及びジャーナルコピーの対象となるディスクアレイド装置200を指定する情報を含んだジャーナルコピー要求コマンドを発行する（ステップ6300。図3の5400）。ジャーナルコピー要求を受取ったセカンダリディスクアレイド装置200Bは、指定されたプライマリデ

ディスクアレイ装置 2 0 0 A に対し、指定された論理ブロック領域に格納されたジャーナルを要求するリードコマンドを発行する。

【 0 0 9 5 】

なお、ジャーナル領域は、図 4 で示すようにメタデータ領域とジャーナルデータ領域に分かれるので、ジャーナルコピー要求ではまずメタデータ領域が指定される。ジャーナルコピー要求を受け取ったセカンダリディスクアレイ装置 2 0 0 B は、メタデータ領域の指定された論理ブロック領域をコピーするリードコマンドを発行し、読み出したメタデータから対応するジャーナルデータが格納されたジャーナルデータ領域の論理ブロック領域を決定する。その後、セカンダリディスクアレイ装置 2 0 0 B は、その論理ブロック領域（すなわち対応するジャーナルデータ）をコピーするリードコマンドを発行する。

【 0 0 9 6 】

又は、セカンダリディスクアレイ装置 2 0 0 がメタデータとジャーナルデータを同時に読み出すリードコマンドを発行しても良い。その場合、各リードコマンドのアドレスとデータ長はポインタから計算される。たとえば、ジャーナルデータについては、正ジャーナルボリュームのジャーナルデータ領域内のリード予定 LBA+1 に対応する L B A からジャーナルイン LBA-1 までの領域が、副ジャーナルボリュームのジャーナルデータ領域内の対応する領域にコピーされる（ステップ 6 4 0 0。図 3 の 5 5 0 0）。

【 0 0 9 7 】

一方、セカンダリホスト 1 0 0 B は、定期的にセカンダリディスクアレイ装置 2 0 0 B のジャーナル処理状況を、ジャーナル作成状況取得コマンドを用いて取得する（ステップ 6 3 1 0、6 3 2 0、6 4 1 0）。具体的には、リード済み LBA 4 2 4 3 及びリード予定 LBA 4 2 4 4 ポインタの値をセカンダリディスクアレイ装置 2 0 0 B から取得する。セカンダリホスト 1 0 0 B は、リード済み LBA 4 2 4 3 及びリード予定 LBA 4 2 4 4 の値が一致した場合、ジャーナルコピー処理（すなわちリード）が完了していると判断する。

【 0 0 9 8 】

尚、リード予定 LBA 4 2 4 4 の情報がセカンダリホスト 1 0 0 B に保持される

場合、セカンダリHOST 100Bは、リード済みLBA 4243を定期的にセカンダリディスクアレイ装置200Bから取得すれば、ジャーナルリード処理の完了を判断することが出来る。

【0099】

ジャーナルコピー処理の完了を確認したセカンダリHOST 100Bは、ジャーナルのリードが完了した副ジャーナルボリュームの論理ブロック領域に対してリストア処理を指示するジャーナルリストア要求コマンドをセカンダリディスクアレイ装置200Bに発行する。尚、副ジャーナルボリュームの容量が大きい場合や、特にユーザがデータのリストアを要求しない場合、その他、早急なリストアが要求されない場合には、リストア処理はジャーナルの転送完了後すぐに行われる必要は無く、別途、ユーザの指示等に基づいて行われる（ステップ6330。図3のステップ5700）。

【0100】

ジャーナルリストア要求コマンドを受取ったセカンダリディスクアレイ装置200Bは、指定されたLBAに対応する論理ブロックに格納されたジャーナルデータを用いて、SVOLに対するリストアを実行する（ステップ6420、図3のステップ5800）。

【0101】

又、ジャーナルコピー処理の終了を確認したセカンダリHOST 100Bは、リード済みLBA 4243で示されるLBAをプライマリHOST 100Aに通知する（ステップ6340）。尚、ジャーナルコピー処理の終了を検出したセカンダリHOST 100Bは、セカンダリディスクアレイ装置200Bに対して次のジャーナルコピー要求を指示することが可能となる（ステップ6350）。

【0102】

リード済みLBA 4243で示されるLBAを通知されたプライマリHOST 100Aは、通知されたLBAに対応する領域までジャーナルをパージするよう、プライマリディスクアレイ装置200Aに指示する（ステップ6120）。指示を受けたプライマリディスクアレイ装置200Aは、指示されたLBAまで、ジャーナルボリュームをパージする（ステップ6210）。

【0103】

図7は、本発明を適用したデータ処理システム50の第二の実施形態を示す図である。説明の便宜のため、第一の実施形態と同じ構成要素については、同じ参照番号を用いている。

【0104】

図7のデータ処理システム50は、ジャーナルコピー処理において、プライマリディスクアレイ装置200Aがセカンダリディスクアレイ装置200Bからのリードコマンドを待つのではなく、プライマリディスクアレイ装置200Aからセカンダリディスクアレイ装置200Bに対してデータを書き込むライトコマンドを発行する点で第一の実施形態と相違する。また、セカンダリディスクアレイ装置200Bの記憶制御装置210がジャーナルリストア処理を実行するのではなく、セカンダリホスト100Bが副ジャーナルボリュームからリストアに使用するジャーナルを読み出して、SVOLのデータをリストアする（図7のステップ5900）点も相違する。本実施形態において、ジャーナルリストアプログラムはセカンダリホスト100B上で実行されるプログラムである。

【0105】

尚、プライマリストレージサブシステム10の構成要素は、セカンダリストレージシステム20のそれらと区別するためにプライマリデバイスまたはプライマリ構成要素と記述されるか、参照番号の後に「A」付きで記述されるか、その両方である（たとえば、プライマリホスト100、ホスト100A、プライマリホスト100A）。同様に、セカンダリデバイスの構成要素はセカンダリデバイスまたはセカンダリ構成要素と記述されるか、参照番号の後に「B」付きで記述されるか、その両方である（たとえば、セカンダリホスト100、ホスト100B、セカンダリホスト100B）。

【0106】

本実施形態においては、ジャーナルコピー処理の主体がプライマリディスクアレイ装置200Aであること、ジャーナルリストア処理を行うのがセカンダリホスト100Bであることから、セカンダリディスクアレイ装置200Bには、特殊な機能を有しない一般的な記憶装置を用いることができる。またデータ処理シ

ステム50はヘテロジニアスなストレージサブシステム、すなわち異なるベンダで製造された、または異なるストレージプロトコルや方法を用いたストレージサブシステムを採用可能である。

【0107】

プライマリサイト10において、PVOLデータの更新（ステップ5100）に対するジャーナル取得処理（ステップ5200）の処理は第一の実施形態と同様である。プライマリホスト100Aは、プライマリディスクアレイ装置200Aから、ジャーナル作成状況に関する情報を随時取得する（ステップ5300）。

【0108】

またプライマリホスト100Aは、プライマリディスクアレイ装置200Aに、セカンダリディスクアレイ装置200Bへのジャーナルコピー要求コマンドを発行する（ステップ5450）。

【0109】

尚、ジャーナルコピー要求コマンドには、ディスクアレイ装置200Bへ送信すべきジャーナルが格納されているジャーナルボリューム、ディスクアレイ装置200Bを指定する情報、及びそのジャーナルをディスクアレイ装置200Bで格納すべきジャーナルボリュームを指定する情報等が含まれている。

【0110】

ジャーナルコピー要求コマンドを受取ったプライマリディスクアレイ装置200Aは、ライトコマンドをセカンダリディスクアレイ装置200Bに発行することで、指定されたジャーナルをセカンダリディスクアレイ装置200Bに送信する（ステップ5600）。セカンダリディスクアレイ装置200Bは、プライマリディスクアレイ装置200Aからのライトコマンドとして受信したジャーナルを、そのコマンドで指定された副ジャーナルボリュームの領域に格納する。

【0111】

その後、非同期コピーマネージャ150を実行するセカンダリホスト100Bは、副ジャーナルボリュームからジャーナルを読み出してSVOL2214へデータのリストアを行なう（ステップ5900）。

【0112】

副ジャーナルボリュームのポインタ管理はセカンダリサイト 2 0 のセカンダリ
 ホスト 1 0 0 B が行ない、ジャーナルコピー要求 5 4 5 0 の作成に必要な情報(
 例えばコピーアドレス計算に必要なリストア済み LBA 4 2 4 1 など)をプライマリ
 ホスト 1 0 0 A に通知する。尚、リストアが終わったジャーナルが格納されてい
 た領域はパーズされ、新たなジャーナルの格納に利用される。

【 0 1 1 3 】

又、本実施形態では、プライマリホスト 1 0 0 A の指示により、プライマリデ
 ィスクアレイ装置 2 0 0 A が、PVOL 2 2 1 2 のイニシャルデータを順にセカンダ
 リディスクアレイ装置 2 0 0 B の SVOL 2 2 1 4 へ書き込む(ライトコマンドをセ
 カンダリディスクアレイ装置 2 0 0 B に発行する)ことで形成コピーが実現する
 。

【 0 1 1 4 】

プライマリホスト 1 0 0 A の指示に基づいて PVOL のすべてのイニシャルデータ
 をセカンダリディスクアレイ装置 2 0 0 B へ書き込んだら、コピープログラム 2
 1 1 0 を実行するプライマリディスクアレイ装置 2 0 0 A は、プライマリホスト
 1 0 0 A へ形成コピーの終了を報告する。プライマリホスト 1 0 0 A は、この報
 告を受信する。これ以降、セカンダリサイト 2 0 でリストアされた SVOL 2 2 1 4
 は、PVOL 2 2 1 2 の内容を反映するボリュームとして扱える。

【 0 1 1 5 】

図 8 は、本発明を適用したデータ処理システムの第三の実施形態を示す図であ
 る。データ処理システム 5 0 は、ホスト間の第一の通信リンク 4 8 を持つが、デ
 ィスクアレイ装置 2 0 0 A と 2 0 0 B 間の第二の通信リンクを持たない。外部記
 憶装置が第二の通信リンクの代わりに用いられる。

【 0 1 1 6 】

本実施形態では、テープ装置またはそれに準ずる第一の外部記憶装置 6 0 がプ
 ライマリホスト 1 0 0 A に、第二の外部記憶装置 6 2 がセカンダリホスト 1 0 0
 B に、それぞれファイバチャネルを介して接続される。外部記憶装置 6 0 と 6 2
 の間はファイバチャネルなどにより接続しても良いし、それら外部記憶装置が磁
 気テープ等の可搬記憶媒体である場合には、必要に応じて装置間で記憶媒体を運

んでデータの移動を行なっても良い。

【0117】

本実施形態においては、プライマリディスクアレイ装置200Aは、第二の実施形態と同様に、PVOL2112についてジャーナル取得処理を行なう。ジャーナルのコピーは、以下の(1)～(3)のステップで行われる。

(1) プライマリホスト100Aが第一の外部記憶装置60に対してデータの書き込みを行なう。具体的には、プライマリホスト100Aは、ジャーナル取得処理開始後、GUIなどの入力装置を介したユーザの指示の入力、または予め定められたスケジュールに応じて正ジャーナルボリュームからジャーナルを読み出し、それを外部記憶装置60に格納する。(ステップ5620)。

(2) 第一の外部記憶装置60に書き込まれたデータを、第二の外部記憶装置62へ移す。本処理は、プライマリホスト100Aとセカンダリホスト100Bのどちらが行なっても良い。データ転送の指示には、例えばANSI (American National Standards Institute) SCSI-3のExtended Copyコマンドが用いられる。

【0118】

データ転送に必要なアドレス情報や、データ転送終了報告などはホスト100Aと100Bの間の通信リンクで交換される。又、可搬記憶媒体を第一の外部記憶装置から第二の外部記憶装置に物理的に移送した後、ユーザや管理者が各ホスト100にデータを移したことを報告しても良い(ステップ5622)。

【0119】

(3) 外部記憶装置62に格納されたデータは、セカンダリホスト100Bの指示に従って、セカンダリディスクアレイ装置200Bに転送される。具体的には、セカンダリホスト100Bは、外部記憶装置62に対してリードコマンド5505を発行し、外部記憶装置62からジャーナルを読み出す。その後、セカンダリホスト100Bは、読み出したジャーナルをもとに、第二の実施形態と同様にSVOL2214のデータをリストアする(ステップ5625)。

【0120】

以上の手順により、PVOLからSVOLへジャーナルボリュームを介したデータ複製

を行なうことができる。尚、形成コピーも同様に行われる。実施形態によっては、外部記憶装置62に格納されたジャーナルはリストア処理が終了しても、特に指示が無い限り消去されない。また外部記憶装置62には、形成コピーの結果、すなわちイニシャルデータも格納されている。

【0121】

さらに、ジャーナルのメタデータ内にPVOL更新時刻のタイムスタンプが含まれているので、本発明のデータ処理システムにおいては、ジャーナル取得処理開始以後任意の時刻におけるPVOLの内容を、セカンダリディスクアレイ装置200BのSVOL2214にリストアすることできる。すなわち、PVOLの形成コピーが完了しているSVOLに対し、セカンダリホスト100Bから指定された時刻以前のタイムスタンプを持つジャーナルを時間順にSVOLへ全てリストアすれば、指定時刻のPVOLの内容をSVOLに再現できる。これを、ポイント・イン・タイム・リカバリと呼ぶ。

【0122】

又、セカンダリディスクアレイ装置200Bのユーザが指定した任意のボリューム2216に対して、ポイント・イン・タイム・リカバリを行うことも可能である。すなわち、外部記憶装置62に格納されているPVOLの形成コピーの結果をまずボリューム2216にコピーし、その後、ボリューム2216へセカンダリホスト100Bから指定された時刻以前の更新時刻のタイムスタンプを持つジャーナルデータを時間順にSVOLへ全てリストアする。

【0123】

なお、ジャーナルデータのリストアでは、同一の領域に対するジャーナルデータが複数存在する場合は、最新のタイムスタンプを持つジャーナルデータのみを使ってリストアする形式としても良い。

又、本実施形態において外部記憶装置60と62は同一の種類の記憶装置でも良いし別々の種類の記憶装置でも良い。またそれらは別の装置として説明したが、それらを同一の装置としても良い。

尚、第一及び第二の実施形態でも、副ジャーナルボリュームに格納されたジャーナルのうち指定した時刻以前のタイムスタンプを持つ全てのジャーナルについ

てSVOLへリストアを行なうことでポイント・イン・タイム・リカバリが実現される。しかしリストアできるPVOLの内容は、副ジャーナルボリュームに格納されている最も古いジャーナルが示すPVOL変更時刻以後のPVOLの内容に限られる。

【0124】

図9は、本発明を適用したデータ処理システムの第四の実施形態を示す図である。本実施形態は第三の実施形態に類似するが、外部記憶装置60及び62がプライマリディスクアレイ装置200A及びセカンダリディスクアレイ装置200Bにそれぞれファイバチャネルで接続される点が異なる。そのため、外部記憶装置50への形成コピーやジャーナルの格納は、プライマリホスト100Aからの指示5450に従って、プライマリディスクアレイ装置200Aが行なう（ステップ5630）。

【0125】

外部記憶装置60に格納されたデータは、プライマリディスクアレイ装置200Aの指示5631に基づくデータ転送または記憶媒体の移動により外部記憶装置62に移される（ステップ5632）。

【0126】

外部記憶装置62からの形成コピーやジャーナルの読み出しは、セカンダリホスト100Bからの指示5400に基づき、セカンダリディスクアレイ装置200Bがリードコマンド5507を発行することで行なわれる（ステップ5635）。ジャーナル取得やリストアの処理は、第一の実施形態に従う。

【0127】

本実施形態によっても、PVOLのデータを、ジャーナルを転送することでSVOLに非同期に複製することができる。本実施形態では、第三の実施形態と異なり、実際のデータ転送をディスクアレイ装置200Aと200Bが行なうので、ホスト100Aと100Bの負荷は低くなる。また、本実施形態でも、ポイント・イン・タイム・リカバリを実現することが出来る。

【0128】

図11は、本発明を適用した計算機システムの第五の実施形態を示す図である。本実施形態は、先述したこれまでの実施形態と異なり、プライマリストレージ

システム10は、複数のセカンダリストレージシステム20及び30と接続される。

【0129】

本実施形態では、プライマリディスクアレイ装置200AのPVOL2212に対応するジャーナルを、セカンダリディスクアレイ装置200BのSVOL2214Bに対応する副ジャーナルボリューム並びにセカンダリディスクアレイ装置200CのSVOL2214Cに対応する副ジャーナルボリュームに各々転送しリストアする。又、PVOL2212から、SVOL2214B並びにSVOL2214Cに対し各々形成コピーも行なう。具体的には、個々のセカンダリディスクアレイ装置200B及び200Cが、プライマリディスクアレイ装置200Aにリードコマンドを発行する。又は、プライマリディスクアレイ装置200Aが、各セカンダリディスクアレイ装置200B、Cにライトコマンドを発行する。これにより、プライマリサイトに格納されたデータの複製が複数のサイトに作成できる。

【0130】

図12は、本発明を適用した計算機システムの第六の実施形態を示す図である。本実施形態においては、ユーザあるいはシステム管理者は、障害などによりセカンダリサイト20が利用できなくなった場合に備え、予め、プライマリホスト100Aに、セカンダリサイト20の代わりに利用できるサイトの候補を一つあるいは複数登録しておく。こうした候補のリストまたはテーブル160がホスト100A内に格納されても良い。

【0131】

セカンダリサイト20が利用できなくなった場合、プライマリホスト100Aは、リスト160から、新たなセカンダリサイト40を選択する。新たなセカンダリサイトの選択の際には、プライマリホスト100Aは、予め定められた優先順位に沿ってセカンダリサイトの選択を行っても良い。この場合、優先順位は、サイトの候補を登録する際に、ユーザが設定する場合が考えられる。あるいは、プライマリサイト100Aと登録されたサイトとの距離、データ転送速度等の条件に基づいてプライマリサイト100Aが自動的に登録されたサイトの選択の優先順位を設定することも考えられる。

【0132】

その後、プライマリホスト100Aは、新たに選択したセカンダリサイト40におけるセカンダリホスト100Dに対し、デバイスグループ等の情報を転送する。デバイスグループ等の情報を受取った新たなセカンダリホスト100Dは、セカンダリホスト100Dに接続されたディスクアレイ装置200Dに対して、新たなSVOL、ジャーナルペアの設定、及び100Dとプライマリホスト100Aとの間のジャーナルコピーを要求する。尚、ほとんどの場合、形成コピーも必要となるので、セカンダリホスト100Dは、形成コピーもディスクアレイ装置200Dに要求する。これらの処理により、リモートレプリケーションの前または最中にセカンダリサイト20で障害が発生した場合にも、新たに選択されたセカンダリサイト40において、プライマリサイト10に格納されたデータの複製作業が継続できる。

【0133】

尚、利用不能になったのがセカンダリサイト20のセカンダリホスト100Bのみで、セカンダリディスクアレイ装置200Bが利用可能である場合は、ディスクアレイ装置200Bをセカンダリサイト40のディスクアレイ装置として引き続き（たとえばストレージエリアネットワークシステム内で）利用することも可能である。

【0134】

図13は、本発明を適用した計算機システムの第七の実施形態を示す図である。

【0135】

本実施形態は、先述したこれまでの実施形態と異なり、プライマリサイト10がプライマリホスト100A及び仮想ディスクアレイ装置15Aから構成され、セカンダリサイト20がセカンダリホスト100B及び仮想ディスクアレイ装置15Bから構成される。各ホスト100は、各仮想ディスクアレイ装置15を単一のディスクアレイ装置200として扱う。すなわち、各ホストは、第一の実施形態と同様のコマンドを仮想ディスクアレイ装置15に発行する。

【0136】

仮想ディスクアレイ装置 1 5 は、バーチャリゼーションサーバ 3 0 0 及び複数のストレージサブシステムから構成される。尚、ここでは、ストレージサブシステムの例として、ディスクアレイ装置 2 0 0 を考える。また、バーチャリゼーションサーバ 3 0 0 は、CPU等を有する一般の計算機である。バーチャリゼーションサーバ 3 0 0 は、プライマリホスト 1 0 0 A（セカンダリホスト 1 0 0 B）、複数の各ディスクアレイ装置 2 0 0 及び他のバーチャリゼーションサーバ 3 0 0 とファイバチャネルで接続される。このファイバチャネルは、第一の実施形態のファイバチャネル 6 6 及び 6 8 に相当し、記憶制御装置 2 1 0 間の通信や、形成コピー及びジャーナルコピー処理に用いられる。また、第一の実施形態と同様にバーチャリゼーションサーバ 3 0 0 間の距離が長い場合にエクステンダ装置を介して ATM などの広域回線を経由してバーチャリゼーションサーバ 3 0 0 間を接続しても良い。

【 0 1 3 7 】

バーチャリゼーションサーバ 3 0 0 は、バーチャリゼーションサーバ 3 0 0 に接続される複数のディスクアレイ装置 2 0 0 が有するボリューム（論理ボリュームでも物理ボリュームでも良い）の集合を、各ホスト 1 0 0 に対し仮想的な一つ（又は複数）の記憶装置システムとして提供する。これはバーチャリゼーションサーバ 3 0 0 が、バーチャリゼーションマネージャと称するプログラム 3 1 0、すなわち接続されたホスト 1 0 0 毎に、接続された各ディスクアレイ装置 2 0 0 上の複数のボリュームを 1 つのアドレス空間を有する記憶領域（以下「仮想ストレージイメージ」）として変換する変換プログラムを実行することで実現する。

【 0 1 3 8 】

ここで、ホスト 1 0 0 と仮想ディスクアレイ装置 1 5 との間のデータ転送について簡単に説明する。たとえば、ホスト 1 0 0 A から仮想ディスクアレイ装置 1 5 A の仮想ストレージイメージに書き込み要求 5 1 0 0 があった場合を考える。この書き込み要求 5 1 0 0 は、バーチャリゼーションサーバ 3 0 0 A によって、ホスト 1 0 0 A に対応する仮想ストレージイメージを構成する各ディスクアレイ装置 2 0 0 への書き込み要求 5 1 0 5 に変換される。その後、バーチャリゼーションサーバ 3 0 0 A は、変換した書き込み要求 5 1 0 5 を、各ディスクアレイ装

置 200 へ送る。このとき、書き込み要求 5100 に含まれる書き込みデータは、上記変換に従って、各ディスクアレイ装置 200 ごとのデータに分割される。また、書き込みアドレスも各ディスクアレイ装置 200 への書き込みアドレスに変換される。

【0139】

次に、ホスト 100B から仮想ディスクアレイ装置 15B の仮想ストレージイメージにデータの読み出し要求があった場合を考える。データ読み出し要求は、バーチャリゼーションサーバ 300B によって、ホスト 100B に対応する仮想ストレージイメージを構成する各ディスクアレイ装置 200 への読み出し要求に変換される。その後、バーチャリゼーションサーバ 300B は、変換した読み出し要求を、各ディスクアレイ装置 200 へ送る。

【0140】

その後、各ディスクアレイ装置 200 からバーチャリゼーションサーバ 300B へ、要求されたデータが転送される（ステップ 5115）。データを受取ったバーチャリゼーションサーバ 300B は、受信したデータを統合し、ホスト 100B に転送する（ステップ 5110）。

【0141】

又、図 13 には示していないが、各バーチャリゼーションサーバ 300 は、各ホスト 100 や各ディスクアレイ装置 200 と同様に IP ネットワークにてリモートコンソールに接続される。ユーザは、リモートコンソールを介してこの計算機システムを管理する。

【0142】

又、例えばバーチャリゼーションサーバ 300 が、バーチャリゼーションサーバ 300 に接続された各ディスクアレイ装置 200 の入出力処理を監視することで、以下の処理を自動的に実現することが考えられる。

(A) 修復できるリードエラー（すなわち、リードデータにエラーが検出されるが、データと共に格納されたエラー訂正符号によって修復される）が頻発するようになったディスクアレイ装置 200 を、別のディスクアレイ装置 200 に置き換えるよう、マッピングを変更する。

(B) アクセス頻度の高いデータを、より高速なディスクアレイ装置200内に配置しなおす。

【0143】

これらの処理に先立って本発明の技術を用いれば、バーチャリゼーションサーバ300の制御に基づくジャーナル取得、ジャーナルコピー、及びジャーナルリストア処理により、予め置き換える元のディスクアレイ装置200上のデータを、置き換える先のディスクアレイ装置200上に複製することができる。その後、仮想ストレージイメージの構成を変更すれば、ホストで実行されるアプリケーションの中断無しでストレージサブシステムの追加や削除が可能である。

【0144】

仮想ディスクアレイ装置15のバーチャリゼーションサーバ300は、ジャーナルバックアップ・リストアプログラム及びコピープログラムを実行する。

【0145】

又、仮想ディスクアレイ装置15は、先の実施形態で説明したようなPVOL、正副ジャーナルボリューム、又はSVOLを有する。ただし、PVOL、正副ジャーナルボリューム、並びにSVOLは、それぞれ複数のディスクアレイ装置200にまたがるように構成されうるが、バーチャリゼーションサーバ300によって、ホスト100、又はバーチャリゼーションサーバ300で実行されるジャーナルバックアップ・リストアプログラム及びコピープログラムからは、第一の実施形態のように（仮想的な）1つのボリュームとして扱われる。従って、バーチャリゼーションサーバ300は、各ホスト100の指示に従って、第一の実施形態と同様の制御、即ち、ジャーナル取得、ジャーナルコピー、ジャーナルリストア、及びジャーナルボリュームの管理を行う。

【0146】

尚、ユーザからの要求や予め決められた方法に応じて、仮想化されるディスクアレイ装置200の数を動的に増減しても良い。更に、プライマリサイト10とセカンダリサイト20で接続されるディスクアレイ装置200が同数・同一種である必要は無い。本実施形態においてバーチャリゼーションサーバ300と各ホスト100は別の装置として説明したが、バーチャリゼーションマネージャをホス

ト100上で動作するプログラムとすると、ホスト100とバーチャリゼーションサーバ300を同一の装置とすることもできる。たとえば、ホスト100がバーチャリゼーションマネージャを実行しても良い。

【0147】

図14は、本発明を適用した計算機システムの第八の実施形態について示す図である。本実施形態は第七の実施形態と類似した仮想ディスクアレイ装置15を用いるが、ジャーナル取得・リストア及びジャーナルコピーは、バーチャリゼーションサーバ300ではなく各ディスクアレイ装置200上のプログラムによって実行される点が先の実施形態とは異なる。

【0148】

又、プライマリサイト10とセカンダリサイト20の各ディスクアレイ装置200は、ファイバチャネルを用いて接続され、ストレージエリアネットワーク（以下SAN）を構成する。すなわち、本実施形態において通信リンク68はSANである。

【0149】

更に、本実施形態では、プライマリサイト10が有するディスクアレイ装置200は、自身が有するボリュームが、セカンダリサイト20のどのディスクアレイ装置200に対応するか、すなわち通信相手となるディスクアレイ装置200がどれかの情報を持たなければならない。またその逆も必要である。そのため、バーチャリゼーションサーバ300は、お互いに各サイトのアドレスマッピング情報を共有し（ステップ3000）、アドレスマッピング情報の変更も共有する。このマッピング情報は各ディスクアレイ装置200に通知される。

【0150】

本実施形態は第七の実施形態と比較して、ジャーナルバックアップ・リストア処理を各ディスクアレイ装置200が行なうので、バーチャリゼーションサーバ300の負荷が低いこと、またプライマリサイト10とセカンダリサイト20間のデータ転送がSANによって行われるので転送速度が速いことが利点である。

【0151】

図16は、第七及び第八の実施形態における、ホスト100と、そのホスト1

00に対応づけられる仮想ストレージイメージを構成する各ディスクアレイ装置200のアドレスマッピングを示すテーブル170を示す図である。図13や図14にはセカンダリホストは100Bの1つしか示していないが、本テーブルでは二つのセカンダリホスト100B、100Cに仮想ストレージイメージを提供する場合を示している。

【0152】

テーブル170には、(1)仮想ストレージイメージを提供する対象のホスト100に関する列172、(2)ホスト100がアクセスする論理ボリューム（以下「ホストアクセスLU」）に関する列174、(3)仮想ストレージイメージを構成するディスクアレイ装置200に関する列176、(4)各ディスクアレイ装置200上の論理ボリューム（以下「ストレージ装置LU」）に関する列178がある。

【0153】

尚、ホスト100とディスクアレイ装置200のアドレスの対応については、図170に示したようなテーブルでなくとも、同様の情報を持つデータ構造、たとえばポインタによるリストを保持することで実現しても良い。

【0154】

上述の実施形態では、形成コピーはジャーナル処理とは別の処理である。しかし、PVOL2212の初期状態を示すジャーナル（以下「基底ジャーナル」）を作成して、ペア生成の後生成された更新ジャーナル（update journal）と共にジャーナル処理することでSVOL2214へPVOL2212の初期データを反映し、形成コピーをジャーナル処理内に包含しても良い。更新ジャーナルは、ペア生成の後、ホストから発行されたライトコマンドに対応するジャーナルである。本発明の説明では、更新ジャーナルは「ジャーナル」または「更新ジャーナル」と呼ぶ。しかし、基底ジャーナルは「基底ジャーナル」としか呼ばない。この用語の区別は、発明の実施の形態でのみ適用し、請求項には適用しない。従って請求項で用いられる「ジャーナル」は基底ジャーナル、更新ジャーナル、マーカジャーナル（後述）、またはそれらの組合わせを示す。

【0155】

具体的には、プライマリホスト100Aが、プライマリディスクアレイ装置200Aに対し基底ジャーナル生成コマンドを発行する。これを受信したプライマリディスクアレイ装置200Aは、PVOL2112の全イニシャルデータから基底ジャーナルを生成する。イニシャルデータは、ペア生成前からPVOL2212に存在したデータである。具体的には、基底ジャーナル生成コマンドを受信したプライマリディスクアレイ装置200Aは、PVOL2212に格納された全てのイニシャルデータを、1つあるいは複数のジャーナルとして正ジャーナルボリュームのジャーナルデータ領域にコピーし、対応するメタデータを各基底ジャーナルに対するメタデータ領域に格納する。ここで、メタデータに含まれるタイムスタンプは、ジャーナルデータがジャーナルデータ領域にコピーされ基底ジャーナルが作成された時刻とする。基底ジャーナルのメタデータにおけるそれ以外の情報（アドレス情報やデータ長など）は更新ジャーナルと同じである。

【0156】

基底ジャーナル作成終了後、プライマリディスクアレイ装置200Aはプライマリホスト100Aに基底ジャーナル生成コマンドの完了を応答する。基底ジャーナルの転送やリストアは、前述の更新ジャーナルに対する処理と同様に行なわれて良い。

【0157】

尚、基底ジャーナルの処理を複数の段階に分け、一回の処理で、イニシャルデータの一部に対して基底ジャーナル作成、ジャーナルコピー、リストアが実行されても良い。例えば正副ジャーナルボリュームの容量がPVOL2212やSVOL2214よりも小さい場合には、まずPVOLに格納されたデータの前半部分の基底ジャーナルが生成され正ジャーナルボリュームに格納されても良い。その後これらの基底ジャーナル（「第一の基底ジャーナル」）が副ジャーナルボリューム2222Bへ転送され、リストアされる。第一の基底ジャーナルの副ジャーナルボリューム2222Bへの転送が終了したら、PVOL2212に格納されたデータの後半部分の基底ジャーナルが生成され、処理される。

【0158】

上述した形成コピーでは、形成コピーとジャーナルリストアが同時に行なわれ

る場合、SVOL2214の各領域のデータが形成コピーのデータでリストアされるかジャーナルボリュームに格納されたデータでリストアされるかを排他的に管理しなければならないが、形成コピーの代わりに基底ジャーナルが生成され処理される場合はそうした管理の必要が無い。

【0159】

また、ジャーナルには、基底ジャーナルと更新ジャーナルに加えてマーカジャーナル (marker journal) が含まれても良い。マーカジャーナルは、特殊なジャーナルで、制御情報をジャーナルコピー処理によってプライマリディスクアレイ装置からセカンダリディスクアレイ装置へ伝えるために用いる。マーカジャーナルは、基底ジャーナルや更新ジャーナルや両方のジャーナルと容易に区別できるように、自身のメタデータ内に目的を示すフラグを持つ。マーカジャーナルは予め定められた条件（たとえば基底ジャーナル生成の完了や中断）でプライマリディスクアレイ装置によって生成され、更新ジャーナルも格納される正ジャーナルボリュームへ格納される。

【0160】

ジャーナルリストア処理の間、セカンダリストレージサブシステムは、リストアするジャーナルがマーカジャーナルであると判断したならば、それをメモリに格納し、予め定められた処理を実行する（たとえば、マーカジャーナルの内容をセカンダリホスト200Bに報告する）。記憶制御装置210がマーカジャーナルを読み、自身のメモリに格納しておいて、ホストから要求を受取ったときにマーカジャーナルの内容をホストへ送信するように実施しても良い。又は、マーカジャーナルの内容のホストへの送信は記憶制御装置210が始めても良い。したがって、マーカジャーナルは、セカンダリストレージサブシステムへ、プライマリストレージサブシステムでのデータ処理に関するイベントの情報（たとえば形成コピーや基底ジャーナル生成の完了、形成コピーや基底ジャーナル生成の中断や再開、ジャーナル取得及びその他の処理）を伝える手段となる。

【0161】

図15は、本発明を適用したデータ処理システムの第九の実施形態について示す図である。本実施形態のデータ処理システムは、第一のサイト10、第二のサ

イト20及び第三のサイト30を有する。第一のサイト10、第二のサイト20及び第三のサイト30は、ネットワーク69（たとえばSAN）で互いに接続されているとする。本システム50では、第一のサイト10と第二のサイト20から構成される第一のサイトグループ8010が定義されているものとする。各サイトはホスト100を有し、そのホスト100は非同期コピーマネージャ150を有する。第一のサイト10が有するホスト100Aは、プライマリホスト100Aとして動作し、第二のサイト20が有するホスト100Bはセカンダリホスト100Bとして動作する。

【0162】

又、其々のサイトのホスト100は、非同期コピーマネージャ150を実行することによって、以下の処理を行う。

【0163】

例えば、あるサイトに異常が発生した場合（例えば、セカンダリホスト100Bに接続されたディスクアレイ装置200Bに障害が発生して使用不可能になった場合）、まず、本データ処理システムは、異常が発生したサイトが、プライマリサイト10か、セカンダリサイト20かを判断する。本例においては、各サイトの各ホスト100が自己に接続された機器の異常の発生を監視する構成としても良い。

【0164】

異常の発生したサイトがセカンダリサイト20であれば、異常を検出したホスト100（ここではセカンダリサイト20のセカンダリホスト100B）が、サイト30のホスト100Cに対して、プライマリサイト10と共に新たなサイトグループを形成するよう、要求する（第二のサイトに障害が発生してもデータの複製が行なえるように）。

【0165】

異常の発生したサイトがプライマリサイト10である場合、異常を発見したサイトは、その異常が発生したプライマリサイト10と組になっているセカンダリサイト20に対して、プライマリサイトに切り替わるよう、要求を行う。要求を受けたセカンダリサイト20は、第三のサイトに対して、セカンダリサイトとな

るよう要求する。

【0166】

尚、プライマリサイト10の障害を発見したサイトが、そのプライマリサイト10と対となるセカンダリサイト20である場合には、セカンダリサイト20は、自サイトをプライマリサイトに切り替える。

【0167】

本実施形態を利用すれば、例えば、障害時の切り替え用サイトとして、世界各地のデータセンタに、本発明に対応することが可能なサイトを構築し、障害復旧までのレンタルサイトとして顧客に貸し出すことが出来る。このサービスに参加する顧客は、顧客が有するサイトのバックアップ用に、さらには、バックアップサイトを使用した際のさらなるバックアップサイトの確保のために、レンタルサイトを使用することが出来る。また、データセンタを運営するサービスプロバイダは、顧客に対して、レンタルサイトが実際に使用されたことに応じて課金することが出来る。サービスプロバイダは、顧客に対するコンフィギュレーション、たとえばレンタルサイトとバックアップサイトの距離、バックアップサイトの数、バックアップサイトの容量や可用性などに応じて、顧客へ課金を行なっても良い。

【0168】

以上、本発明者によってなされた発明を実施例の形態に基づき具体的に説明したが、本発明は前記実施の形態に限定されるものでなく、その要旨を逸脱しない範囲で種々変更可能であることはいうまでもない。

【0169】

上述した本発明のデータ処理システムでは、ディスクアレイ装置がジャーナル取得・リストア及びジャーナルコピー処理を行い、ホスト側でジャーナル管理、リモートコピー状態管理を行わせる構成とする。これにより、プライマリサイトとセカンダリサイト間でのデータの複製は、ホスト間で制御命令をやり取りして行われ、実際のデータ転送はディスクアレイ装置間のファイバケーブル等で実施される。このことによって、ホスト間の一般回線のトラフィックを最小限に抑えることができ、またコピーも専用線のため、より高速な回線となり処理性能の向

上が可能である。

【0170】

さらに、プライマリサイトとセカンダリサイトのデータ移動には、専用線その他、テープ等の外部記憶装置を使用することで、ユーザ指定の任意の時点のジャーナルを外部記憶装置から読み出してジャーナルリストア処理に用いることが可能になる。この機能により、ユーザの必要な時点のデータへの回復も可能とする効果がある。

【0171】

さらに、ディスクアレイ装置がジャーナルをライトコマンドにより他のディスクアレイ装置に書き出す機能を有し、そのデータをホストで読み込み、リストアすることで、セカンダリサイトのディスクアレイに特別な機能を持たせなくてもデータ転送及びデータの複製を実現することが可能となる。

【0172】

【発明の効果】

本発明によって、複数サイト間でのデータ転送又はデータ複製をする際に、ホスト間の一般回線のトラフィックを抑えることができ、またデータ転送性能が向上する。さらに、ユーザの必要な時点のデータの回復もできる。

【0173】

また、多種多様なサイト間でのデータの複製を容易に実行することが出来る。

【0174】

さらに、セカンダリサイトのディスクアレイ装置に特別な機能を持たせなくても良い。つまり、本発明によれば、一般的に互換性の無い例えば製造元の異なるディスクアレイ装置間の接続することが容易に行なえる。

【図面の簡単な説明】

【図1】

本発明を適用したデータ処理システムの第一の実施形態のハードウェア構成を示す図である。

【図2】

図1に示した本発明の第一の実施形態における処理手順を示したフローチャー

トである。

【図 3】

図 1 に示した本発明の第一の実施形態におけるジャーナルの取得、コピー及びリストアの手順を示す図である。

【図 4】

図 1 に示した本発明の第一の実施形態における PVOL と正ジャーナルボリューム 2 2 2 2 A の対応関係を示す図である。

【図 5】

本発明における正副ジャーナルのジャーナルデータ領域の対応関係を示した図である。

【図 6】

本発明の第一の実施形態におけるジャーナル取得、データ転送、ジャーナルリストアの手順の詳細を示すフローチャートである。

【図 7】

本発明を適用したデータ処理システムの第二の実施形態を示す図である。

【図 8】

本発明を適用したデータ処理システムの第三の実施形態を示す図である。

【図 9】

本発明を適用したデータ処理システムの第四の実施形態を示す図である。

【図 1 0】

本発明を適用したデータ処理システムの第一の実施形態の論理構成を示す図である。

【図 1 1】

本発明を適用したデータ処理システムの第五の実施形態を示す図である。

【図 1 2】

本発明を適用したデータ処理システムの第六の実施形態を示す図である。

【図 1 3】

本発明を適用したデータ処理システムの第七の実施形態を示す図である。

【図 1 4】

本発明を適用したデータ処理システムの第八の実施形態を示す図である。

【図 1 5】

本発明を適用したデータ処理システムの第九の実施形態を示す図である。

【図 1 6】

バーチャリゼーションサーバ 3 0 0 B 内のアドレステーブルを示す図である。

【図 1 7】

本発明における副ジャーナルボリュームと SVOL の対応関係を示す図である。

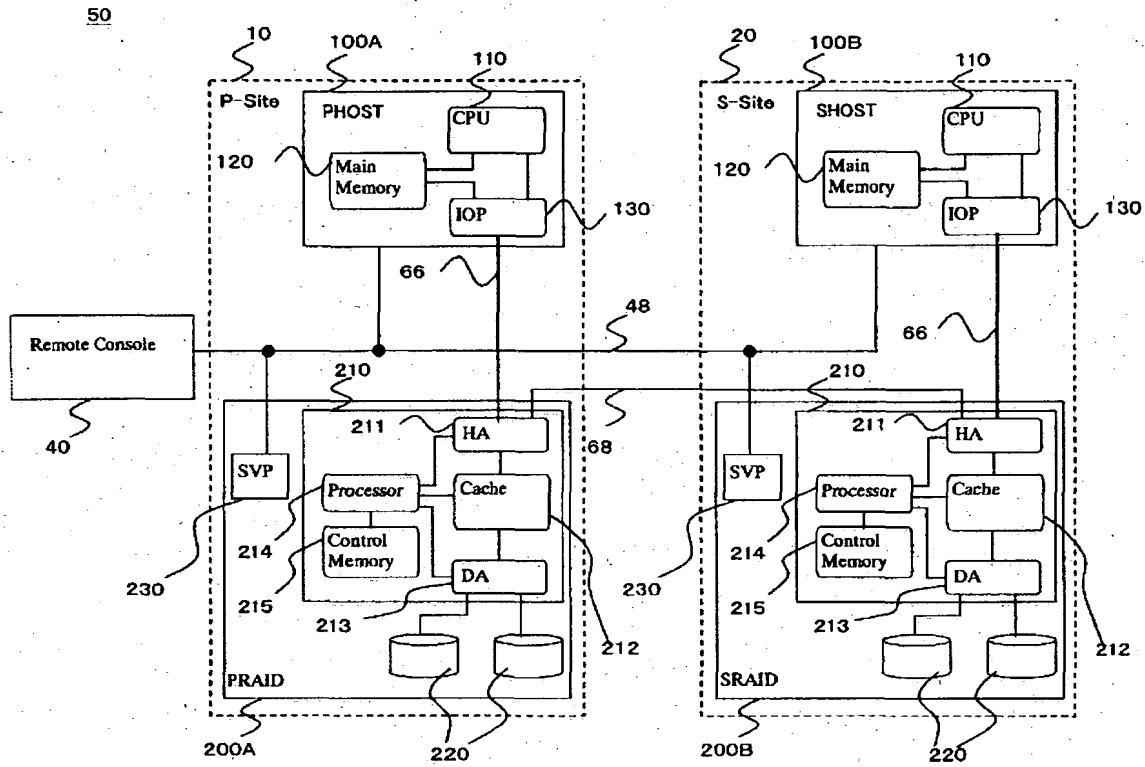
【符号の説明】

1 0 … プライマリサイト、 2 0 … セカンダリサイト、 1 0 0 A … プライマリホ
スト、 2 0 0 A … プライマリディスクアレイ装置、 1 0 0 B … セカンダリホスト
、 2 0 0 B … セカンダリディスクアレイ装置。

【書類名】 図面

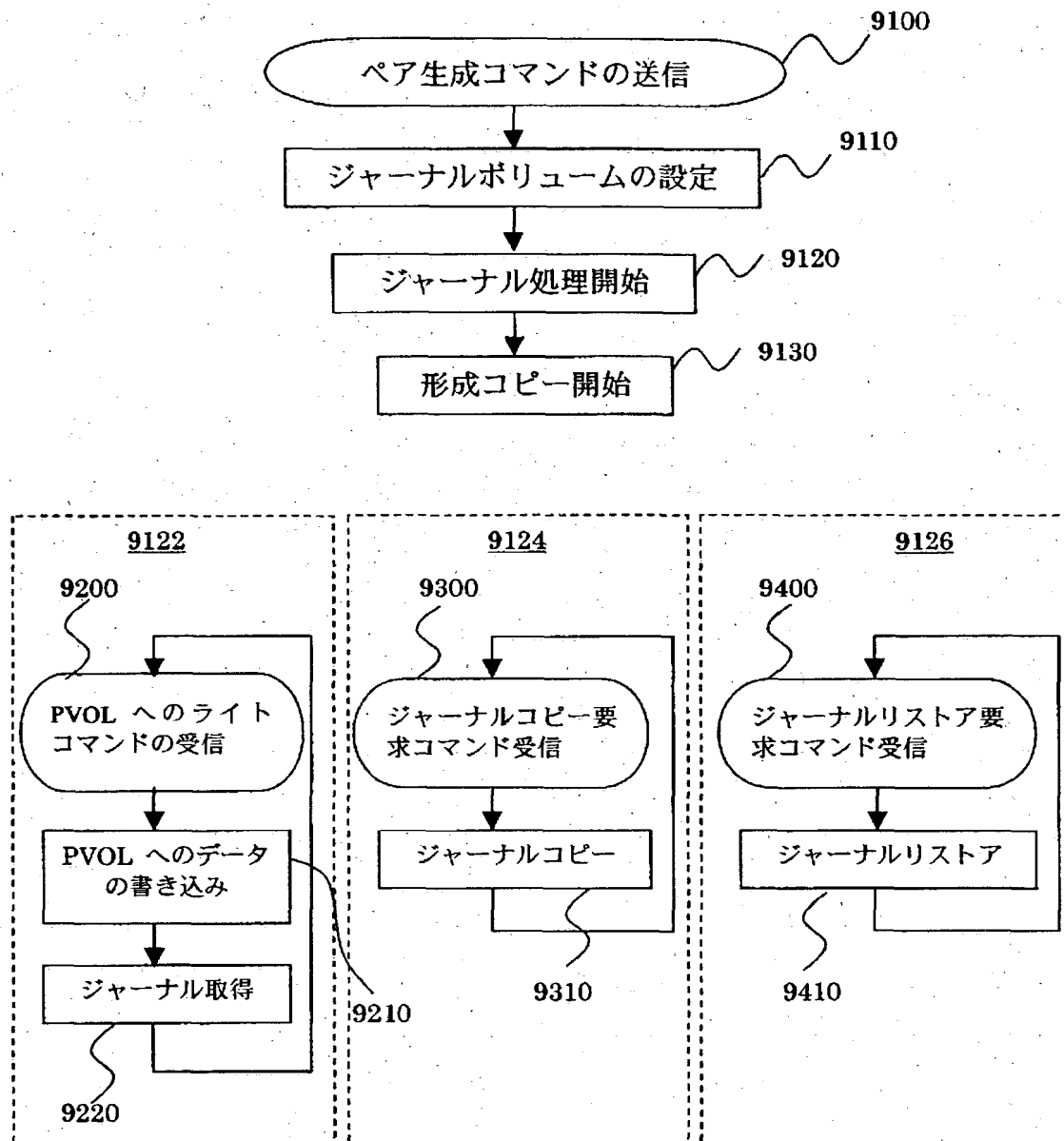
【図1】

図 1



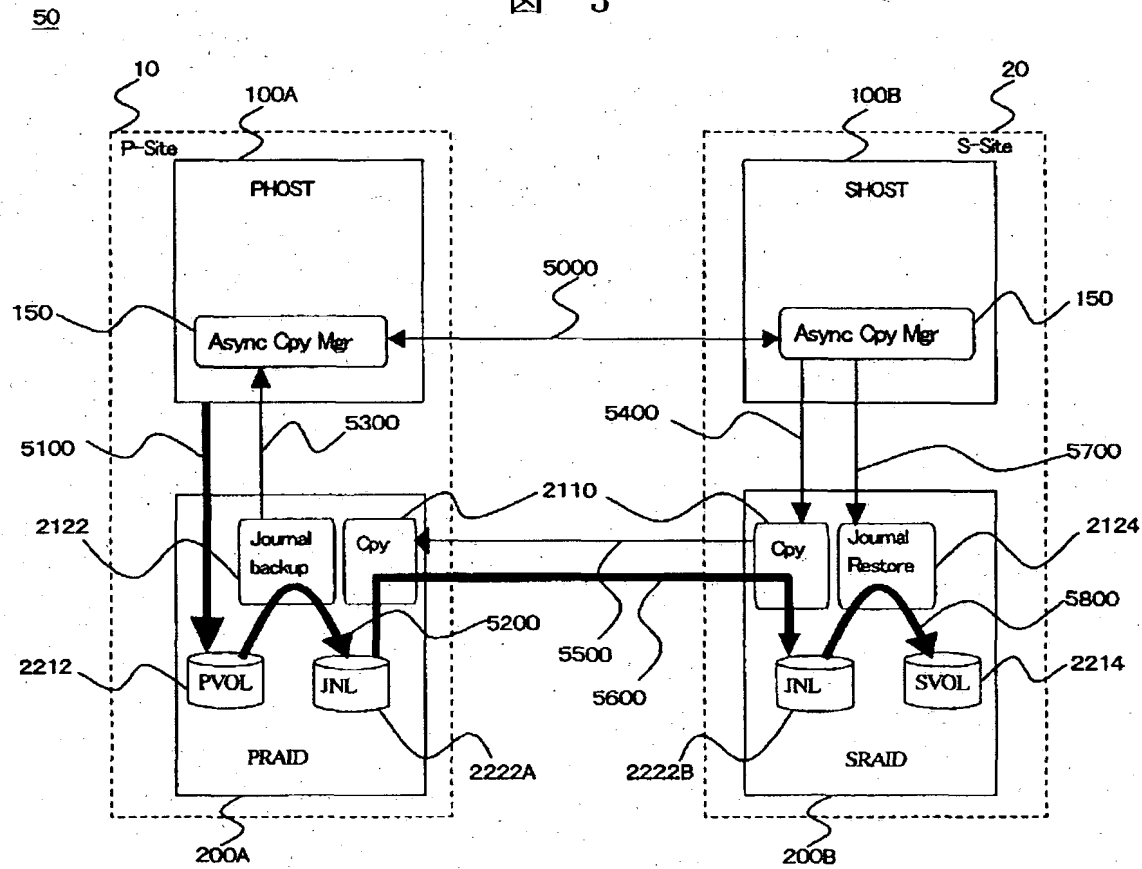
【図 2】

図 2



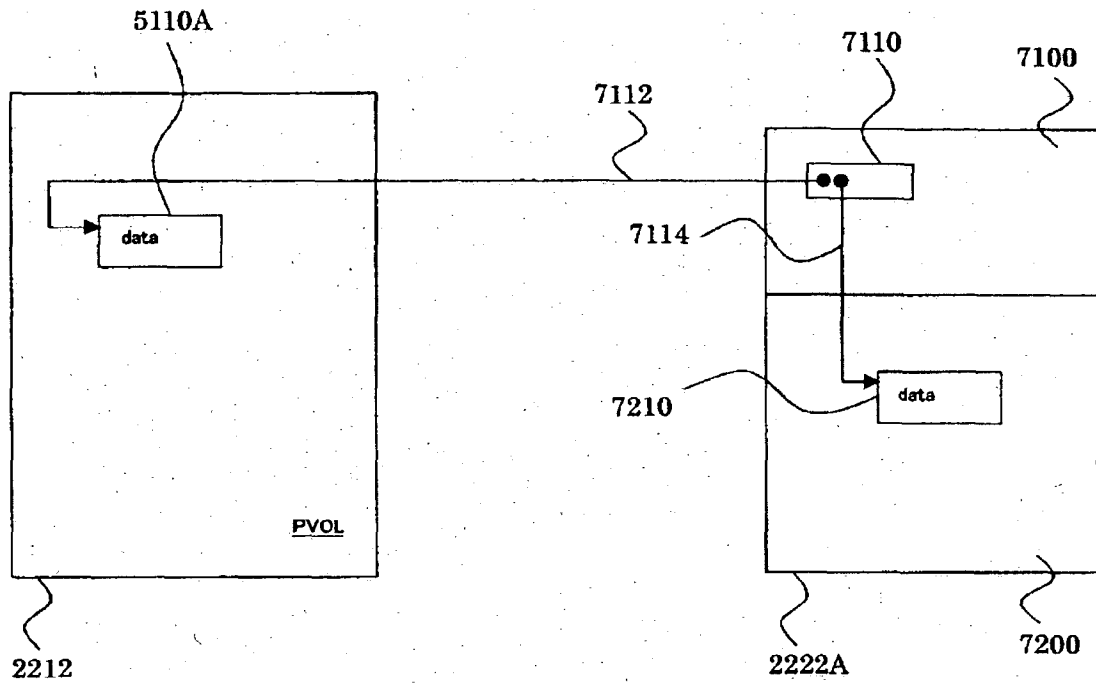
【図 3】

図 3



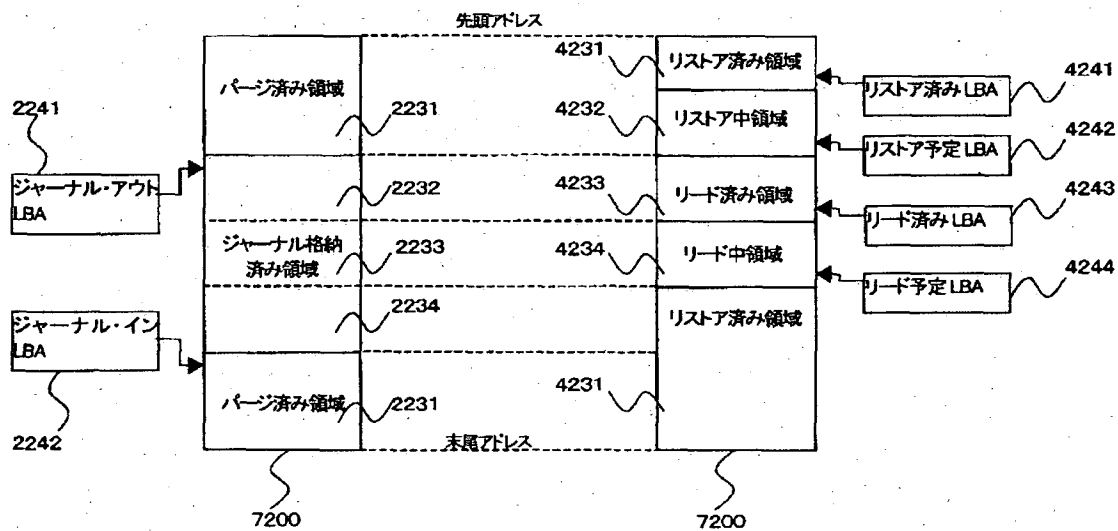
【図 4】

図 4



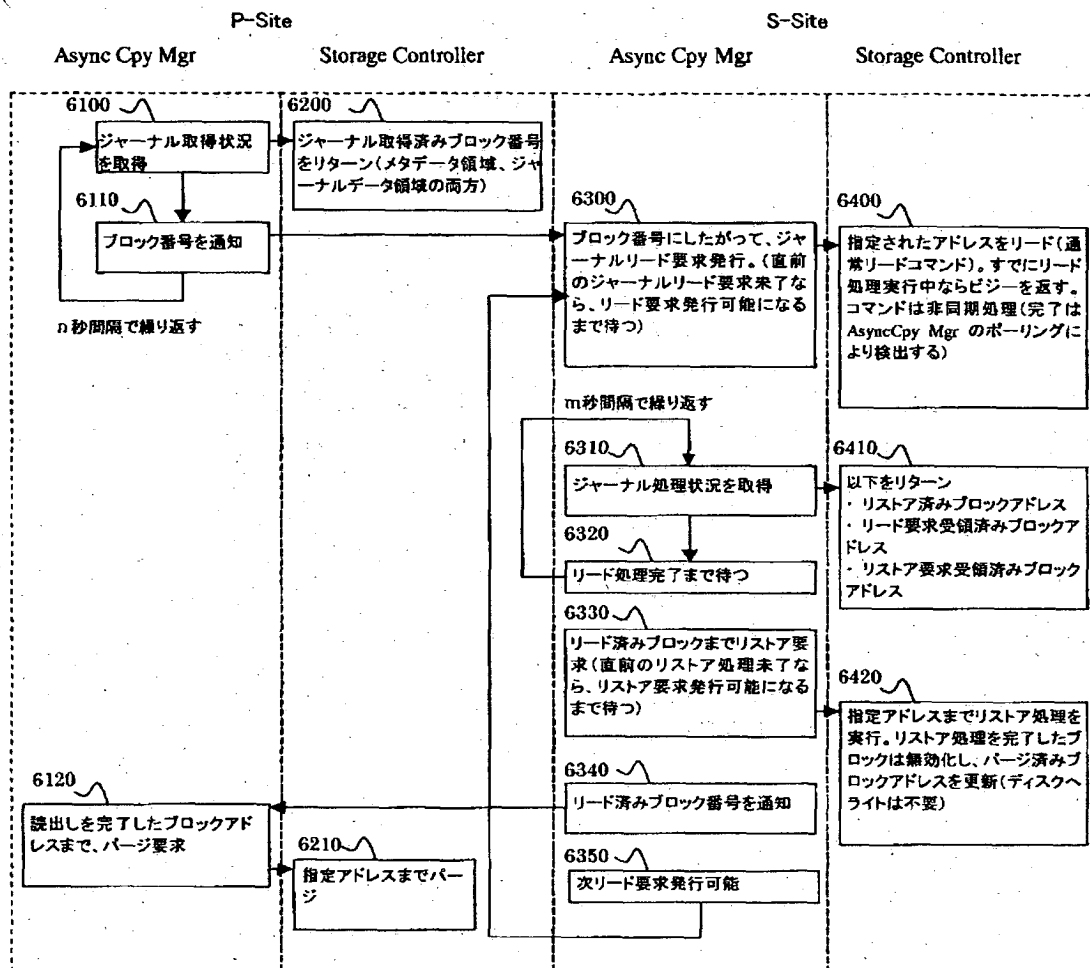
【図 5】

図 5

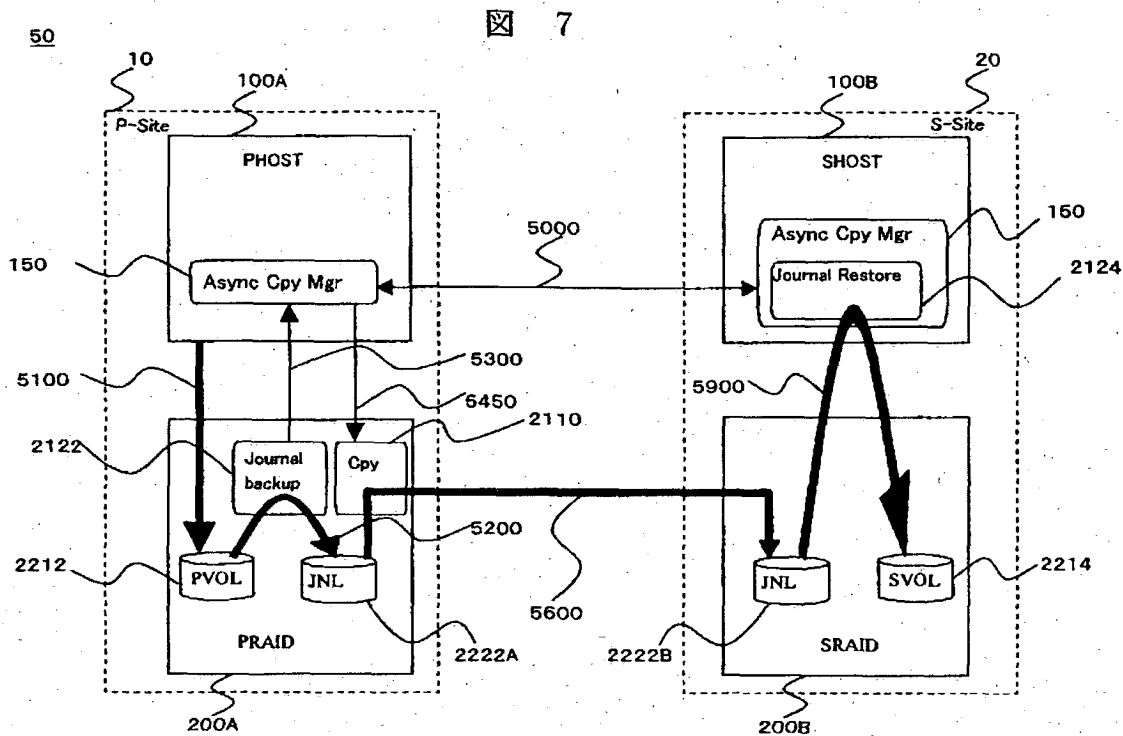


【图 6】

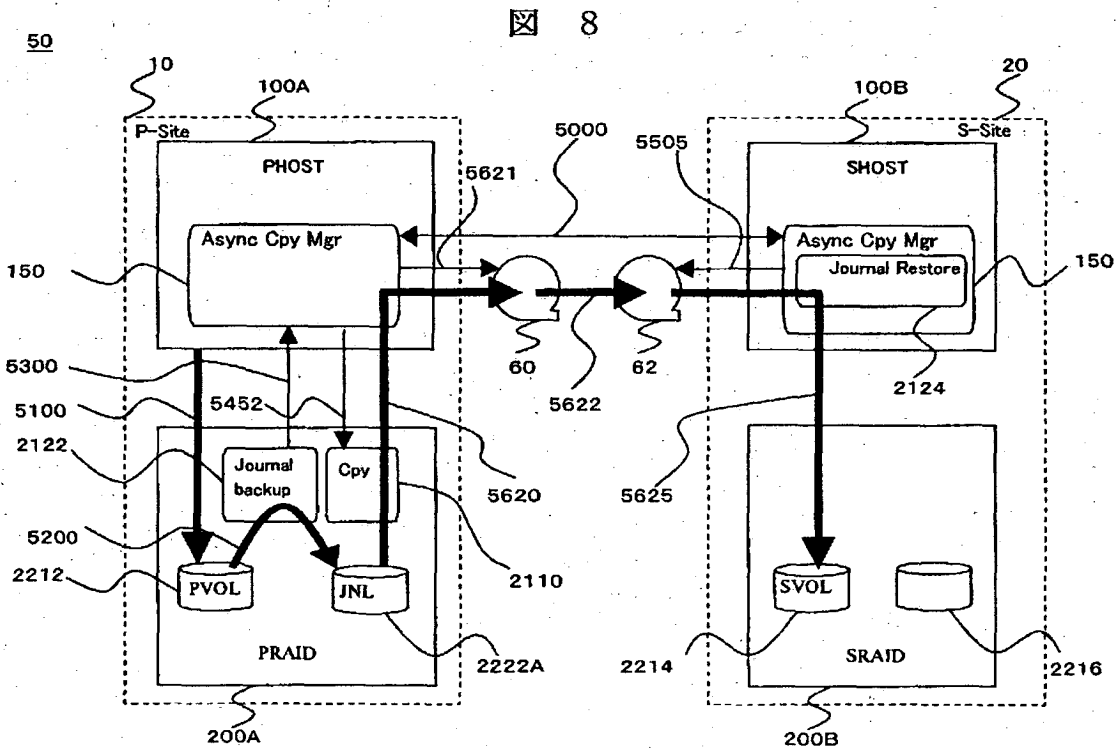
圖 6



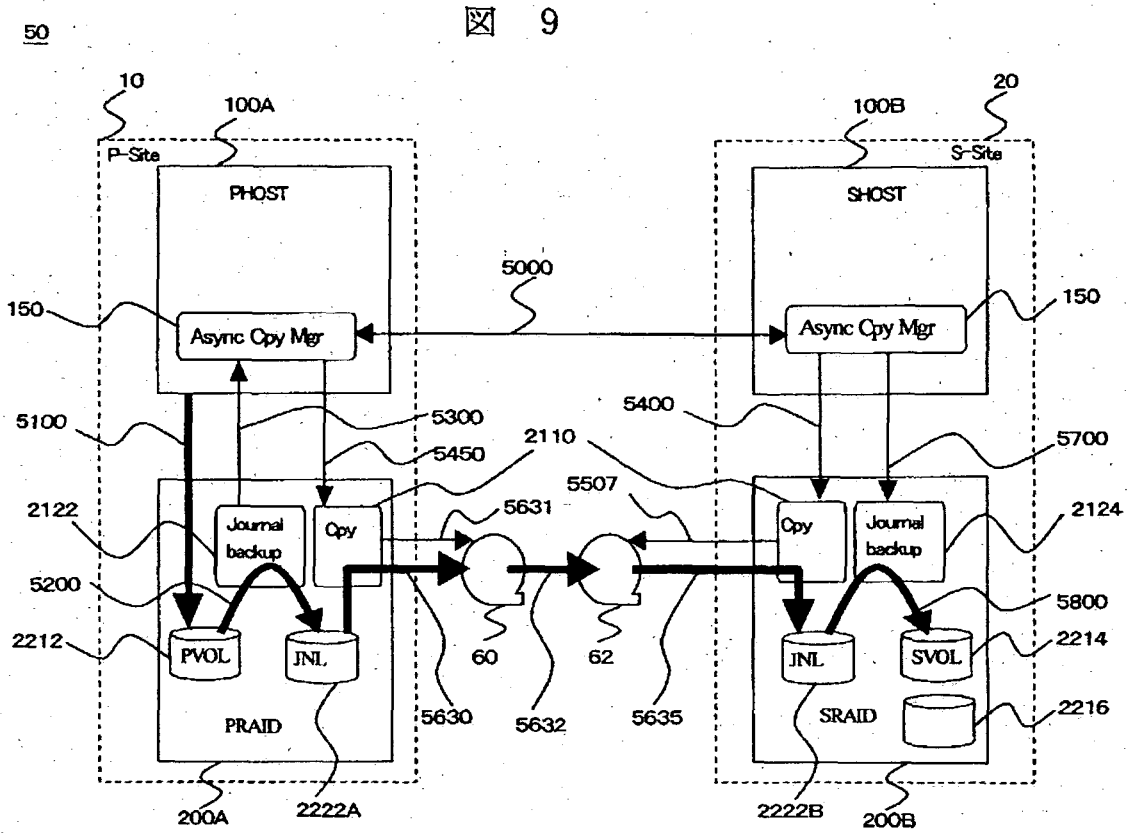
【図7】



【図8】

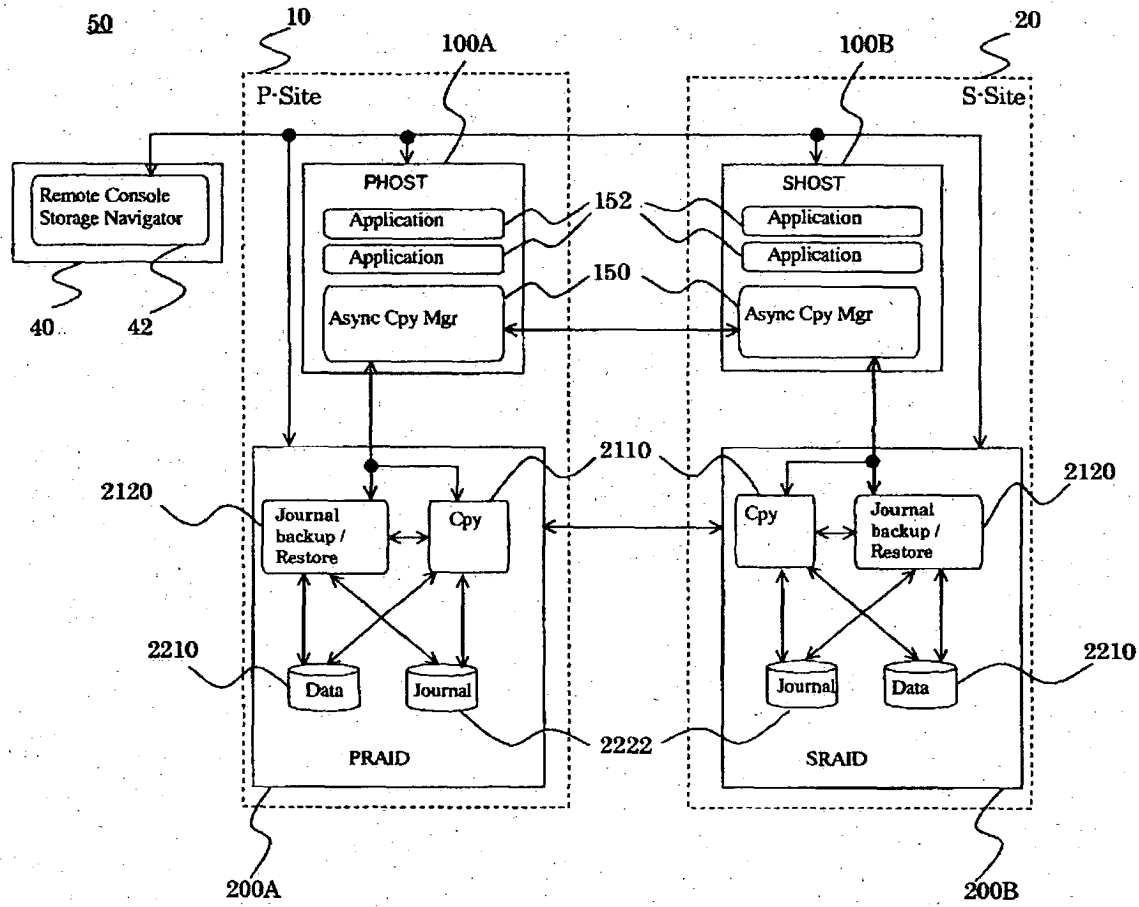


【図9】

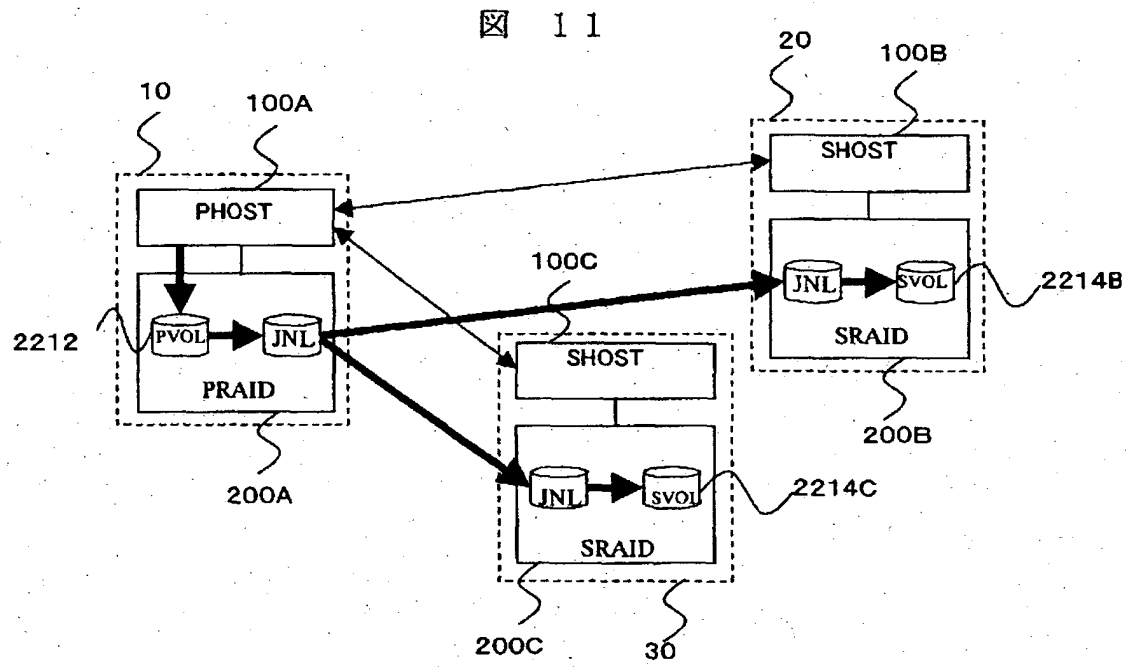


【図10】

図 10

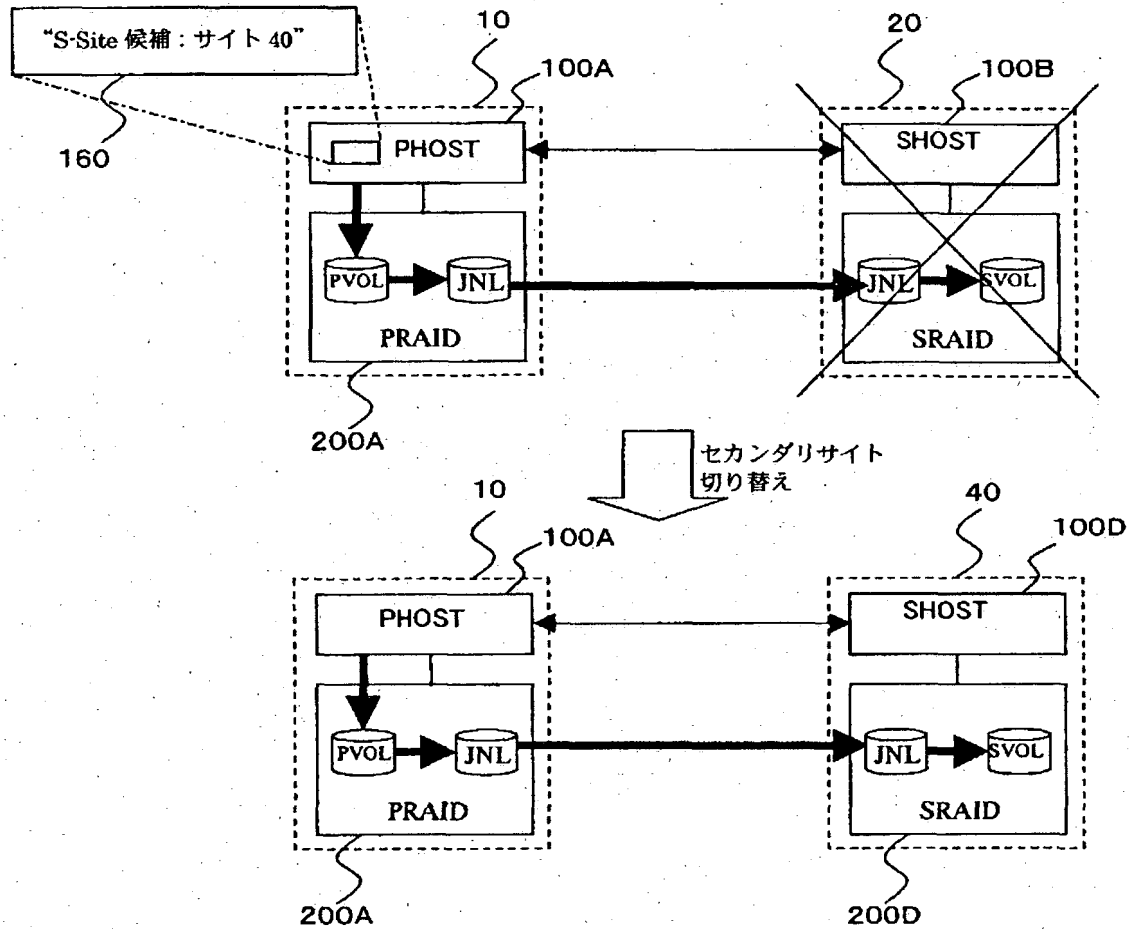


【図11】



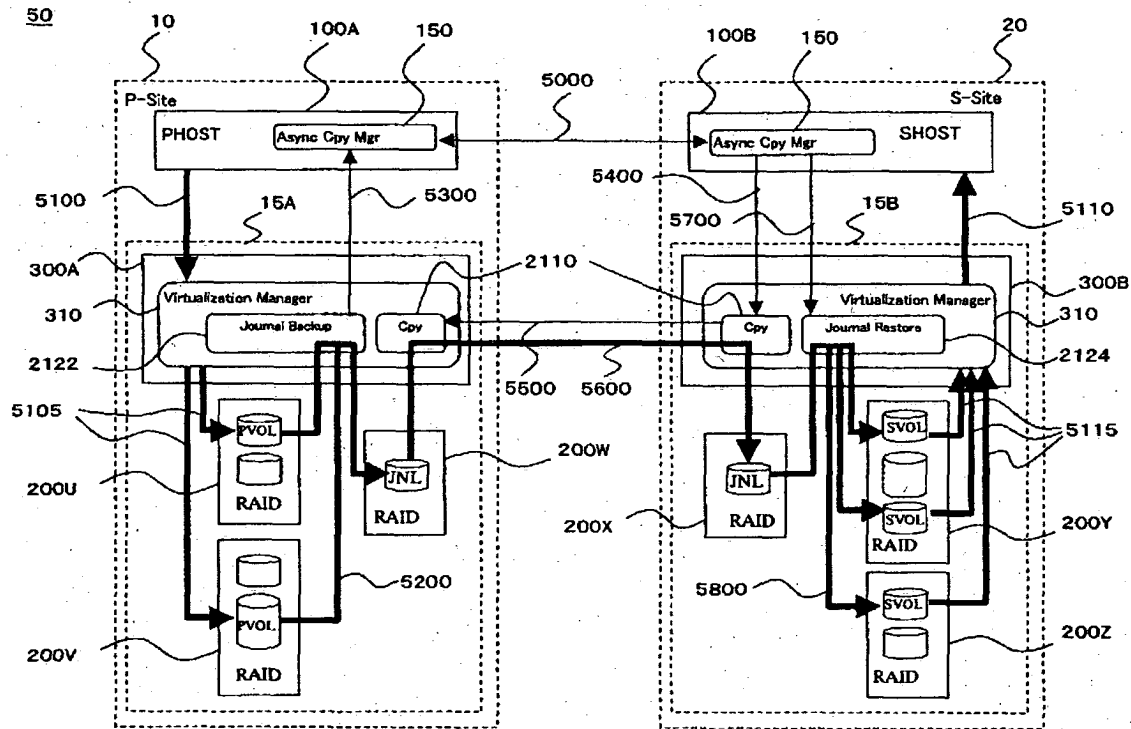
【図12】

図 12



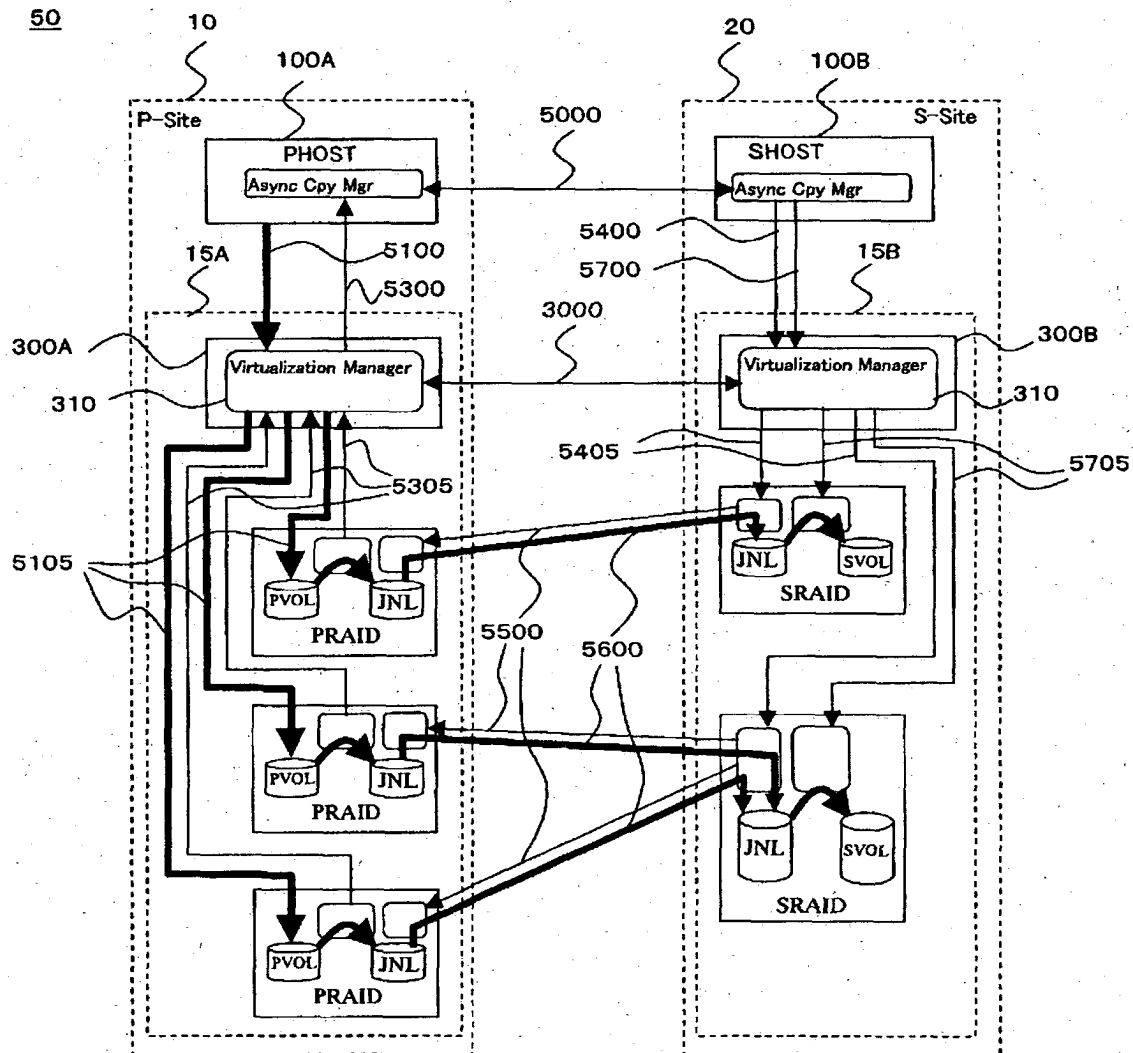
【図13】

図 13



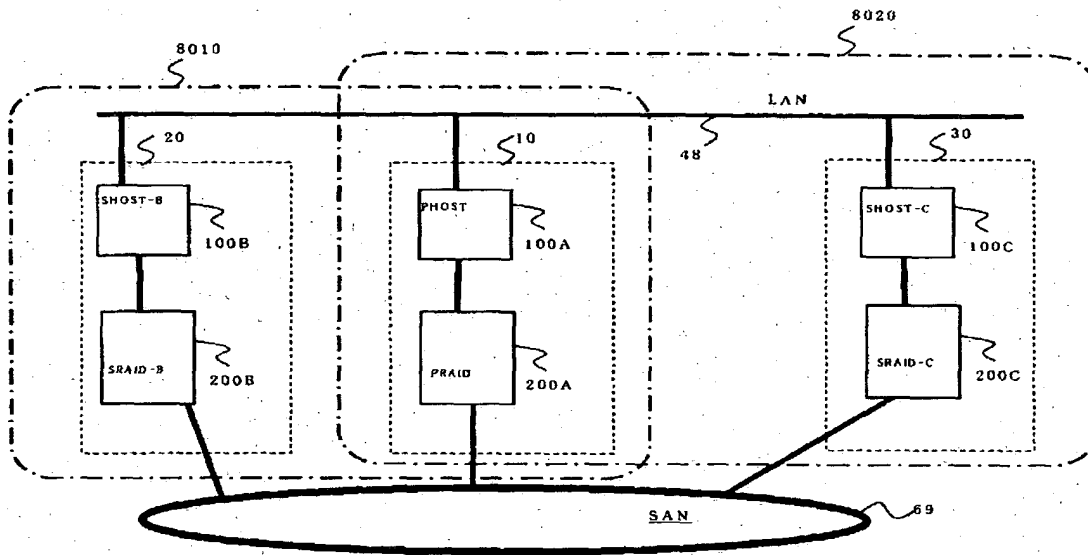
【図14】

図 14



【図 15】

図 15



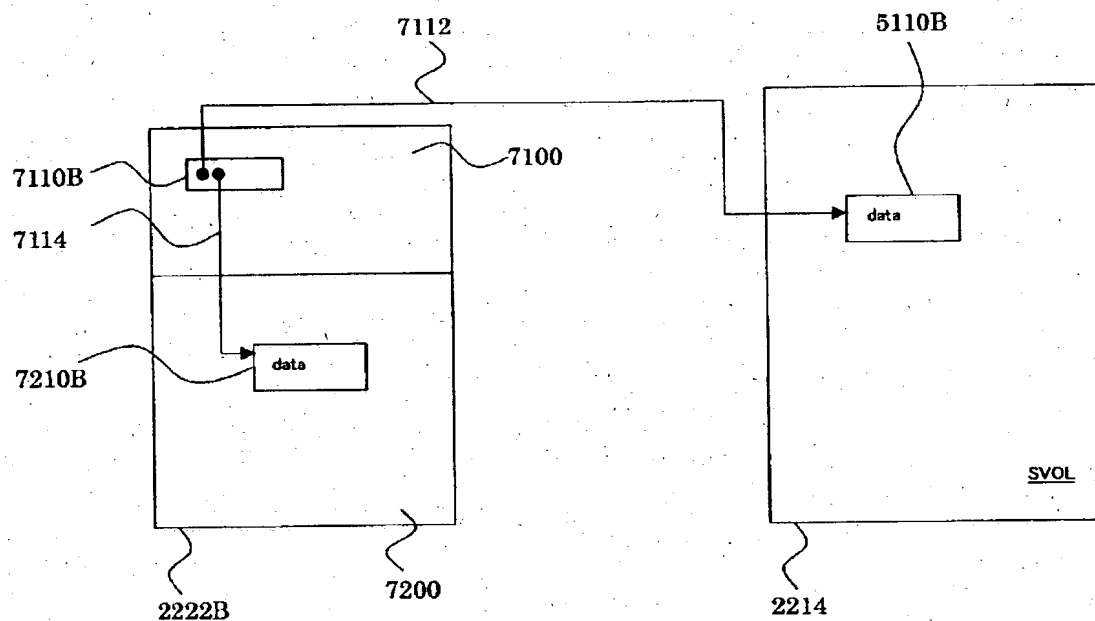
【図 16】

図 16

Host	Host access LU	Storage device	Storage device LU
100B	0 ~ 3	200X	0 ~ 3
	4 ~ 7	200Y	0 ~ 3
	8 ~ 11	200Y	12 ~ 15
	12 ~ 15	200Z	16 ~ 19
100C	0 ~ 7	200Y	4 ~ 11
	8 ~ 11	200Z	4 ~ 7

【図 17】

図 17



【書類名】 要約書

【要約】

【課題】

ホストベースリモートサイトへのデータの複製は、ホストに負荷がかかる。

【解決手段】

実際のデータ転送は各サイトが有する記憶装置間で、その制御等はホストで行うことにより、ホストに負荷をかけずに、より複雑なシステムにおいても、リモートサイトへデータを複製することが出来る。また、ジャーナルを用いることにより、ユーザの要求に合ったデータの複製を作成することができる。

【選択図】 図1

特2003-050244

認定・付加情報

特許出願の番号	特願2003-050244
受付番号	50300313952
書類名	特許願
担当官	第七担当上席 0096
作成日	平成15年 2月28日

<認定情報・付加情報>

【提出日】 平成15年 2月27日

出 願 人 履 歴 情 報

識別番号 [000005108]

1. 変更年月日 1990年 8月31日
[変更理由] 新規登録
住 所 東京都千代田区神田駿河台4丁目6番地
氏 名 株式会社日立製作所